GENETICS OF NEURODEGENERATIVE DISEASES

EDITED BY: Rahul Desikan, Celeste Karch and Leo P. Sugrue
PUBLISHED IN: Frontiers in Neuroscience

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88971-780-4
DOI 10.3389/978-2-88971-780-4

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

GENETICS OF NEURODEGENERATIVE DISEASES

Topic Editors:
Rahul Desikan, University of California, San Francisco, United States
Celeste Karch, Washington University in St. Louis, United States
Leo P. Sugrue, University of California, San Francisco, United States

Citation: Desikan, R., Karch, C., Sugrue, L. P., eds. (2021). Genetics of Neurodegenerative Diseases. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88971-780-4

Table of Contents

04 ***Evaluation of Gene-Based Family-Based Methods to Detect Novel Genes Associated With Familial Late Onset Alzheimer Disease***
Maria V. Fernández, John Budde, Jorge L. Del-Aguila, Laura Ibañez, Yuetiva Deming, Oscar Harari, Joanne Norton, John C. Morris, Alison M. Goate, NIA-LOAD family study group, NCRAD and Carlos Cruchaga

23 ***Pleiotropic Effects of Variants in Dementia Genes in Parkinson Disease***
Laura Ibanez, Umber Dube, Albert A. Davis, Maria V. Fernandez, John Budde, Breanna Cooper, Monica Diez-Fairen, Sara Ortega-Cubero, Pau Pastor, Joel S. Perlmutter, Carlos Cruchaga and Bruno A. Benitez

33 ***Combining Polygenic Hazard Score With Volumetric MRI and Cognitive Measures Improves Prediction of Progression From Mild Cognitive Impairment to Alzheimer's Disease***
Karolina Kauppi, Chun Chieh Fan, Linda K. McEvoy, Dominic Holland, Chin Hong Tan, Chi-Hua Chen, Ole A. Andreassen, Rahul S. Desikan, and Anders M. Dale for the Alzheimer's Disease Neuroimaging Initiative

40 ***Plin4-Dependent Lipid Droplets Hamper Neuronal Mitophagy in the MPTP/p-Induced Mouse Model of Parkinson's Disease***
Xiaojuan Han, Jialei Zhu, Xinlei Zhang, Qiqi Song, Jianhua Ding, Ming Lu, Sifan Sun and Gang Hu

54 ***Insulin-Like Growth Factor Binding Protein 2 Is Associated With Biomarkers of Alzheimer's Disease Pathology and Shows Differential Expression in Transgenic Mice***
Luke W. Bonham, Ethan G. Geier, Natasha Z. R. Steele, Dominic Holland, Bruce L. Miller, Anders M. Dale, Rahul S. Desikan, Jennifer S. Yokoyama and for the Alzheimer's Disease Neuroimaging Initiative

64 ***Targeted Sequencing of Alzheimer Disease Genes in African Americans Implicates Novel Risk Variants***
Mark W. Logue, Daniel Lancour, John Farrell, Irina Simkina, M. Daniele Fallin, Kathryn L. Lunetta and Lindsay A. Farrer

75 ***New Antibody-Free Mass Spectrometry-Based Quantification Reveals That C9ORF72 Long Protein Isoform Is Reduced in the Frontal Cortex of Hexanucleotide-Repeat Expansion Carriers***
Arthur Viodé, Clémence Fournier, Agnès Camuzat, François Fenaille, NeuroCEB Brain Bank, Morwena Latouche, Fanny Elahi, Isabelle Le Ber, Christophe Junot, Foudil Lamari, Vincent Anquetil and François Becher

86 ***Microbiome-Derived Lipopolysaccharide (LPS) Selectively Inhibits Neurofilament Light Chain (NF-L) Gene Expression in Human Neuronal-Glial (HNG) Cells in Primary Culture***
Walter J. Lukiw, Lin Cong, Vivian Jaber and Yuhai Zhao

98 ***Genetic Overlap Between Alzheimer's Disease and Bipolar Disorder Implicates the MARK2 and VAC14 Genes***
Ole Kristian Drange, Olav Bjerkehagen Smeland, Alexey A. Shadrin, Per Ivar Finseth, Aree Witoelar, Oleksandr Frei, Psychiatric Genomics Consortium Bipolar Disorder Working Group, Yunpeng Wang, Sahar Hassani, Srdjan Djurovic, Anders M. Dale and Ole A. Andreassen

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 04 April 2018
doi: 10.3389/fnins.2018.00209

Evaluation of Gene-Based Family-Based Methods to Detect Novel Genes Associated With Familial Late Onset Alzheimer Disease

Maria V. Fernández[1,2], John Budde[1,2], Jorge L. Del-Aguila[1,2], Laura Ibañez[1,2], Yuetiva Deming[1,2], Oscar Harari[1,2], Joanne Norton[1,2], John C. Morris[2,3], Alison M. Goate[4], NIA-LOAD family study group[†], NCRAD[†] and Carlos Cruchaga[1,2]*

[1] Department of Psychiatry, Washington University School of Medicine, St. Louis, MO, United States, [2] Hope Center for Neurological Disorders, Washington University School of Medicine, St. Louis, MO, United States, [3] Knight Alzheimer's Disease Research Center, Washington University School of Medicine, St. Louis, MO, United States, [4] Department of Neuroscience, Ronald M. Loeb Center for Alzheimer's Disease, Icahn School of Medicine at Mount Sinai, New York, NY, United States

OPEN ACCESS

Edited by:
Michael F. Miles,
Virginia Commonwealth University,
United States

Reviewed by:
Fanny Elahi,
University of California, San Francisco,
United States
Iris Broce,
University of California, San Francisco,
United States

***Correspondence:**
Carlos Cruchaga
cruchagac@wustl.edu

[†]Membership of the NIA-LOAD and NCRAD Family Study Group is provided in the Acknowledgements

Specialty section:
This article was submitted to Neurogenomics, a section of the journal Frontiers in Neuroscience

Received: 26 December 2017
Accepted: 15 March 2018
Published: 04 April 2018

Citation:
Fernández MV, Budde J, Del-Aguila JL, Ibañez L, Deming Y, Harari O, Norton J, Morris JC, Goate AM, NIA-LOAD family study group, NCRAD and Cruchaga C (2018) Evaluation of Gene-Based Family-Based Methods to Detect Novel Genes Associated With Familial Late Onset Alzheimer Disease. Front. Neurosci. 12:209. doi: 10.3389/fnins.2018.00209

Gene-based tests to study the combined effect of rare variants on a particular phenotype have been widely developed for case-control studies, but their evolution and adaptation for family-based studies, especially studies of complex incomplete families, has been slower. In this study, we have performed a practical examination of all the latest gene-based methods available for family-based study designs using both simulated and real datasets. We examined the performance of several collapsing, variance-component, and transmission disequilibrium tests across eight different software packages and 22 models utilizing a cohort of 285 families (N = 1,235) with late-onset Alzheimer disease (LOAD). After a thorough examination of each of these tests, we propose a methodological approach to identify, with high confidence, genes associated with the tested phenotype and we provide recommendations to select the best software and model for family-based gene-based analyses. Additionally, in our dataset, we identified *PTK2B*, a GWAS candidate gene for sporadic AD, along with six novel genes (*CHRD*, *CLCN2*, *HDLBP*, *CPAMD8*, *NLRP9*, and *MAS1L*) as candidate genes for familial LOAD.

Keywords: gene-based, family-based, clustering, variance-component, transmission disequilibrium, rare variants, whole exome sequencing, Alzheimer's disease

INTRODUCTION

Alzheimer disease (AD) is a complex condition for which almost 50% of its phenotypic variability is due to genetic causes; yet only 30% of the genetic variability is explained by known markers (Ridge et al., 2016). GWAS studies have identified more than 20 risk loci (Lambert et al., 2013) and sequencing studies have identified additional genes harboring low frequency variants with large effect size (*TREM2*, *PDL3*, *UNC5C*, *SORL1*, and *ABCA7*; Sims et al., 2017). Recent studies also indicate that Late-Onset AD (LOAD) families are enriched for genetic risk factors (Cruchaga et al., 2017). Therefore, studying those families may lead to the identification of novel variants and genes (Guerreiro et al., 2013; Cruchaga et al., 2014).

Current consensus is that the missing heritability for complex traits like AD may be hidden within rare variants that have low to moderate effect on disease risk (Frazer et al., 2009; Manolio et al., 2009; Cirulli and Goldstein, 2010). The rarity of these markers requires specific study designs and statistical analyses for their detection. The simplest approach to detect rare variants for association is to test each variant individually using standard contingency table and regression methods. But due to the limited number of observations of the rare minor allele for a specific variant, the statistical power to detect association with any rare variant is limited; hence, extremely large samples are required and a more stringent multiple-test correction is necessary (Li and Leal, 2008; Bansal et al., 2010). It has been acknowledged that the best alternative to single-variant analysis is to collapse sets of pre-defined candidate rare variants within significant units, usually genes (gene-based sets) (Neale and Sham, 2004; Lee et al., 2014). For collapsing tests each variant is given a certain weight and the weights of all variants within the region are summed; depending on the weights and how summation is performed there are three major types of gene-based methods: collapsing tests, variance-component tests, and combined tests (Lee et al., 2014). Collapsing tests analyze whether the overall burden of rare variants is significantly different between cases and controls by regressing disease status on minor allele counts (MAC). The Cohort Allelic Sum Test (CAST) is a dominant genetic model which assumes that the presence of any rare variant increases disease risk (Morgenthaler and Thilly, 2007); whereas the Combined Multivariate and Collapsing (CMC) method collapses rare variants in different MAF categories and evaluates the joint effect of common and rare variants through Hoteling's test (Li and Leal, 2008). However, neither CAST nor CMC tests account for directional effect. The Variable Threshold (VT) test does allow for both trait-increasing and trait-decreasing variants; it selects optimal frequency thresholds for burden tests of rare variants and estimates *p*-values analytically or by permutation (Price et al., 2010). Variance-componence methods test for association by evaluating the distribution of genetic effects for a group of variants while appropriately weighting the contribution of each variant. The sequence kernel association test (SKAT) casts the problem to mixed models (Lee et al., 2014) and, in the absence of covariates, SKAT reduces to a C-alpha test (Neale et al., 2011). Finally, collapsing and variance component tests can be combined into one statistical method, the SKAT-O approach (Lee et al., 2012), which is statistically efficient regardless of the direction and effect of the variants tested.

All these methods were initially designed for unrelated case-control studies; but considering the rarity of these variants, large datasets are required to achieve statistical power (Laird and Lange, 2006). Alternatively, family-based studies in which several family members share the same phenotype may provide more statistical power than regular case-controls studies (Li et al., 2006; Cirulli and Goldstein, 2010; Kazma and Bailey, 2011; Ott et al., 2011). Pioneering methods for gene-based analyses in familial datasets are based on the transmission disequilibrium test (TDT–Spielman et al., 1993) which uses the marker genotype of an affected child and genotypes of the parents to test for association (Laird et al., 2000; Horvath et al., 2001; Ott et al., 2011; De et al., 2013; Ionita-Laza et al., 2013). TDT works under the paradigm of Mendel's laws to determine which marker in the affected offspring is responsible for the phenotype (Zöllner et al., 2004). TDT methods have been extended to test rare-variants by grouping information across multiple variants within a genomic region (He et al., 2014). However, these methods were still not valid for incomplete or nuclear families that have several affected offspring. Considering the late-onset nature of Alzheimer disease it is often difficult to obtain genetic information from parents (to conform trios) or nuclear family units. The typical pedigree in familial LOAD represents incomplete, large familial units (**Figure 1**). Most of the early software for gene-based family-based studies were not suitable for complex pedigrees like those observed in Alzheimer studies. In recent years gene-based methods, whether referring to collapsing, variance-component, or transmission disequilibrium tests, have been adapted to account for complex family structure in its gene-based calculations. Among the software that can manage large pedigrees we find SKAT (Wu et al., 2011), FSKAT (Yan et al., 2015), GSKAT (Wang et al., 2013), RV-GDT (He et al., 2017), EPACTS (http://genome.sph.umich.edu/wiki/EPACTS), FarVAT (Choi et al., 2014), PedGene (Schaid et al., 2013), and RareIBD (Sul et al., 2016).

In this study, we wanted to evaluate the performance of the eight most common gene-based family-based methods available by using a real dataset, over 250 multiplex families affected with Alzheimer disease, under different conditions and models. We simulated multiple scenarios in which candidate variants in the same gene perfectly segregates with disease status to rank the different programs and models. We also tested the performance of these tests for identifying known causal genes for AD in our cohort. Finally, we performed genome-wide analyses to evaluate the power of each of these tests. Altogether, we discuss the pros and cons of each method that can be informative for other investigators performing similar analyses: complex diseases in complex, incomplete, large families. We want to emphasize that although this work focused on AD, the information extracted from this work can be applied to other complex traits as well. Finally, based on the results from the methods analyzed, we present some candidate genes for AD.

MATERIALS AND METHODS

Cohort

The LOAD families included in this study originated from two cohorts: Washington University School of Medicine (WUSM; $n = 1{,}144$) and Alzheimer Disease Sequencing Project (ADSP; $n = 91$) (**Table 1**).

WUSM Cohort

Samples from the Washington University School of Medicine (WUSM) cohort were recruited by either the Charles F. and Joanne Knight Alzheimer's Disease Research Center (Knight ADRC) at the WUSM in Saint Louis or the National Institute on Aging Genetics Initiative for Late-Onset Alzheimer's Disease (NIA-LOAD). This study was approved by each recruiting center's Institutional Review Board and research was carried out

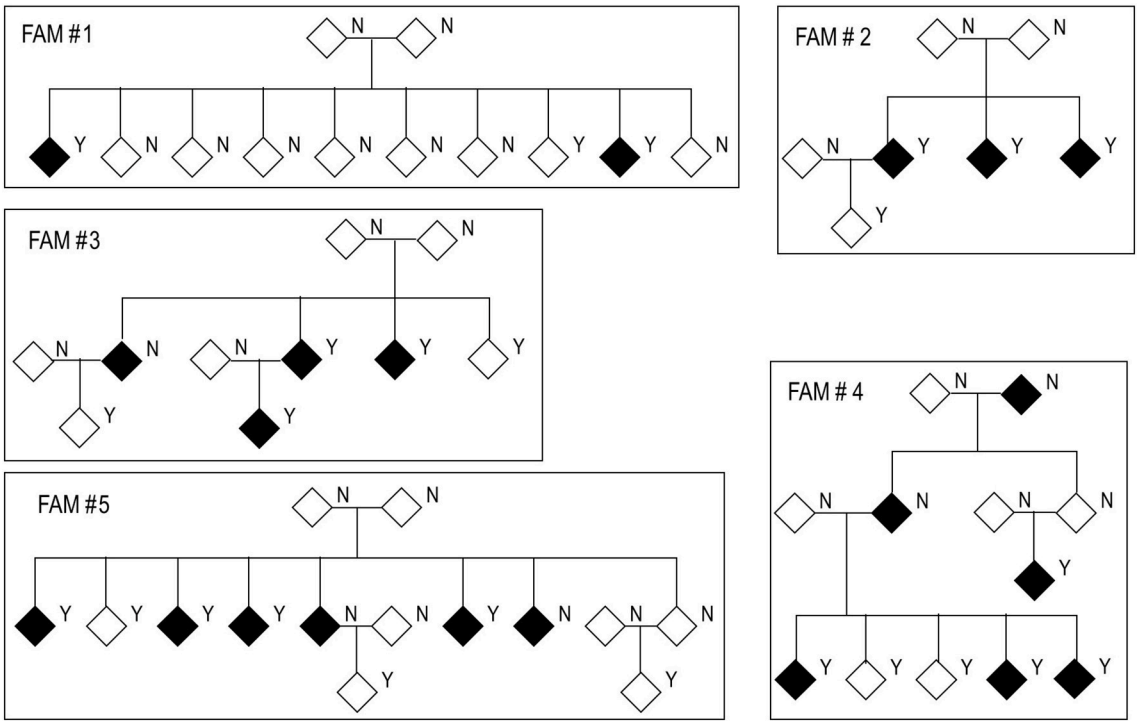

FIGURE 1 | Structure of families used in this study. Black diamonds represent cases and white diamonds represent controls. Y: genetic data available. N: no genetic data available.

TABLE 1 | Demographic data for the familial dataset employed in this study.

	N	*Age ± SD	*Age range	% Fe	% APOE4
Cases	824	73 ±7	48–99	63	73
Controls	411	83 ± 9	39–104	59	51
Total	1235	77 ± 10	39–104	61	65

**Age At Onset (AAO) for cases and Age at Last Assessment (ALA) for controls.*

in accordance with the approved protocol. Written informed consent was obtained from participants and their family members by the Clinical and Genetics Core of the Knight ADRC. The approval number for the Knight ADRC Genetics Core family studies is 201104178. The NIA-LOAD Family Study has recruited multiplex families with two or more siblings diagnosed with LOAD across the United States. A description of these samples has been reported previously (Wijsman et al., 2011; Cruchaga et al., 2012; Fernández et al., 2017). We selected individuals for sequencing from families in which APOEε4 did not segregate with disease status, and in which the proband of the family did not carry any known mutation in *APP*, *PSEN1*, *PSEN2*, *MAPT*, *GRN*, or *C9orf72* (described previously; Cruchaga et al., 2012).

ADSP Cohort

The Alzheimer's Disease Sequencing Project (ADSP) is a collaborative work of five independent groups across the USA that aims to identify new genomic variants contributing to increased risk for LOAD (https://www.niagads.org/adsp/content/home). During the discovery phase, ADSP generated whole genome sequence (WGS) data from members of multiplex LOAD families, and whole exome sequence (WES) data from a large case-control cohort. These data are available to qualified researchers through the database of Genotypes and Phenotypes (https://www.ncbi.nlm.nih.gov/gap Study Accession: phs000572.v7.p4).

The familial cohort of the ADSP consists of 582 individuals from 111 multiplex AD families from European-American, Caribbean Hispanic, and Dutch ancestry (details about the samples are available at NIAGADS). We downloaded raw data (.sra format) from dbGAP for 143 IDs (113 cases and 23 controls) from 37 multiplex families of European-American ancestry that were incorporated with the WUSM cohort.

Sequencing

Samples were sequenced using either whole-genome sequencing (WGS, 12%) or whole-exome sequencing (WES, 88%). Exome libraries were prepared using Agilent's SureSelect Human All Exon kits V3 and V5 or Roche VCRome (**Table 2**). Both WES and WGS samples were sequenced on a HiSeq2000 with paired end reads, with a mean depth of coverage of 50× to 150× for WES and 30× for WGS. Alignment was conducted against GRCh37.p13 genome reference. Variant calling was performed separately for WES and WGS following GATK's 3.6 Best Practices (https://software.broadinstitute.org/gatk/best-practices/) and restricted to Agilent's V5 kit plus a 100 bp of

TABLE 2 | Number of samples for which whole genome sequencing (WGS) or whole exome sequencing (WES) was performed, with detail of the exon library kits employed in this study.

Exon library kit	WGS	WES
WGS	153	
Agilent's SureSelect Human All Exon kits V3	0	28
Agilent's SureSelect Human All Exon kits V5	0	665
Roche VCRome	0	389
Total	153	1,082

padding added to each capture target end. We used BCFTOOLS (https://samtools.github.io/bcftools/bcftools.html) to decompose multiallelic variants into biallelic prior to variant quality control. Variant Quality Score Recalibration (VQSR) was performed separately for WES and WGS, and for SNPs and INDELs. Only those SNPs and indels that fell above the 99.9 confidence threshold, as indicated by WQSR, were considered for analysis; variants within low complexity regions were removed from both WES and WGS and variants with a depth (DP) larger than the average DP + 5 SD in the WGS dataset were removed. At this point SNPs and indels from WES and WGS datasets were merged into one file. Non-polymorphic variants and those outside the expected ratio of allele balance for heterozygosity calls (ABHet = 0.3–0.7) were removed. Additional hard filters implemented included quality depth ($QD \geq 7$ for indels and $QD \geq 2$ for SNPs), mapping quality ($MQ \geq 40$), fisher strand balance ($FS \geq 200$ for indels and $FS \geq 60$ for SNPs), Strand Odds Ratio (SOR ≥ 10 for Indels and SOR ≥ 3 for SNPs), Inbreeding Coefficient (IC ≥ -0.8 for indels) and Rank Sum Test for relative positioning of reference vs. alternative alleles within reads (RPRS ≥ -20 for Indels and RPRS ≥ -8 for SNPs) (**Figure S1**). We used PLINK1.9 (https://www.cog-genomics.org/plink2/ibd) to remove variants that were out of Hardy Weinberg equilibrium ($p < 1 \times 10^{-6}$), with a genotype calling rate below 95%, with differential missingness between cases vs. controls, WES vs. WGS, or among different sequencing platforms ($p < 1 \times 10^{-6}$).

Samples with more than 10% of missing variants (four samples) and whose genotype data indicated a sex discordant from the clinical database (three samples) were removed from the dataset. Individual and familial relatedness was confirmed using identity-by-descent (IBD) calculations, an existing GWAS dataset for these individuals, and the pedigree information. Because many of the ADSP families were also recruited from the NIA-LOAD repository there is a certain overlap (48 individuals) between the WUSM and the ADSP familial cohorts; we kept the duplicate that had better genotyping rate after QC. Principal Component Analysis (PCA) was calculated to corroborate ancestry and restrict our analysis to only samples from European American origin. Functional impact and population frequencies of variants were annotated with SnpEff (Cingolani et al., 2012). For this analysis, only SNVs with a minor allele frequency (MAF) below 1%, as registered in ExAC (Lek et al., 2016), were tested.

We excluded families carrying a known pathogenic mutation in any of the Mendelian genes for Alzheimer disease, Frontotemporal Dementia, or Parkinson disease (Fernández et al., 2017). We restricted the selection of families to those with at least one case and one control in the family, and we excluded any participants that were initially clinically diagnosed with AD but had a different diagnosis after pathological examination. Finally, our dataset consisted of 1,235 non-hispanic whites (NHW), 824 cases and 411 controls, from 285 different families (**Table 1**, **Table S1**). Of these 1,235 individuals, 1144 originated from WUSM and 91 were from ADSP.

Study Design and Analysis

The goal of this study was to test the performance and power of different gene-based family-based methods currently available, using a real dataset consisting of 1,235 non-hispanic white individuals from 285 families densely affected with AD. We created three different scenarios to test (**Figure 2**). First, using the real phenotype and pedigree structure from 25 of the 285 families, we generated a synthetic dataset with multiple variants and families with perfect segregation. Second, we evaluated different variant-combinations for the *APOE* gene. Third, we performed genome-wide gene-based analysis of only nonsynonymous SNPs with a MAF <1%. For each one of these scenarios we evaluated the performance of the different gene-based methods (collapsing, variance-component, and transmission disequilibrium) from the following family-based packages: SKAT (Wu et al., 2011), FSKAT (Yan et al., 2015), GSKAT (Wang et al., 2013), RVGDT (He et al., 2017), EPACTS (http://genome.sph.umich.edu/wiki/EPACTS), FarVAT (Choi et al., 2014), PedGene (Schaid et al., 2013), RareIBD (Sul et al., 2016). Some of these software offer the option to run different gene-based algorithms; e.g., GSKAT, EPACTS, FarVAT or PedGene can run collapsing and variance-component tests; therefore, we ran a total of 25 models (**Table 3**). The details of each one of these scenarios are described next.

Simulated Data

We selected 25 representative families from our entire dataset for which there were genotypic data for three to seven members (**Table S2**). We used the existing family structure and phenotypes of these families, and a simulated gene called "GENE-A" containing five variants. We generated several scenarios in which different numbers of families presented perfect segregation with disease status for a variant in GENE-A (**Table 4**, **Table S2**). First, we considered a scenario in which only the first five families of the dataset were included in the analyses and each family presented a different perfectly segregating variant of GENE-A [scenario 5 family carriers (FC) and 0 non-carriers (FNC): 5FC×0FNC]. Second, we generated additional scenarios in which we kept the same five families as carriers of segregating variants in GENE-A, and added five (scenario 5FC×5FNC), ten (scenario 5FC×10FNC), 15 (scenario 5FC×15FNC), and 20 (scenario 5FC×20FNC) families that were not carriers of any variant in GENE-A. Then, we considered four scenarios of 25 families in which each new scenario added families who were carriers of a segregating variant in GENE-A. We started with the scenario 5FC×20FNC, then we simulated 10 families who had carriers and 15 families who were non-carriers (scenario 10FC×15FNC), 15 families with carriers and 10 families who were non-carriers

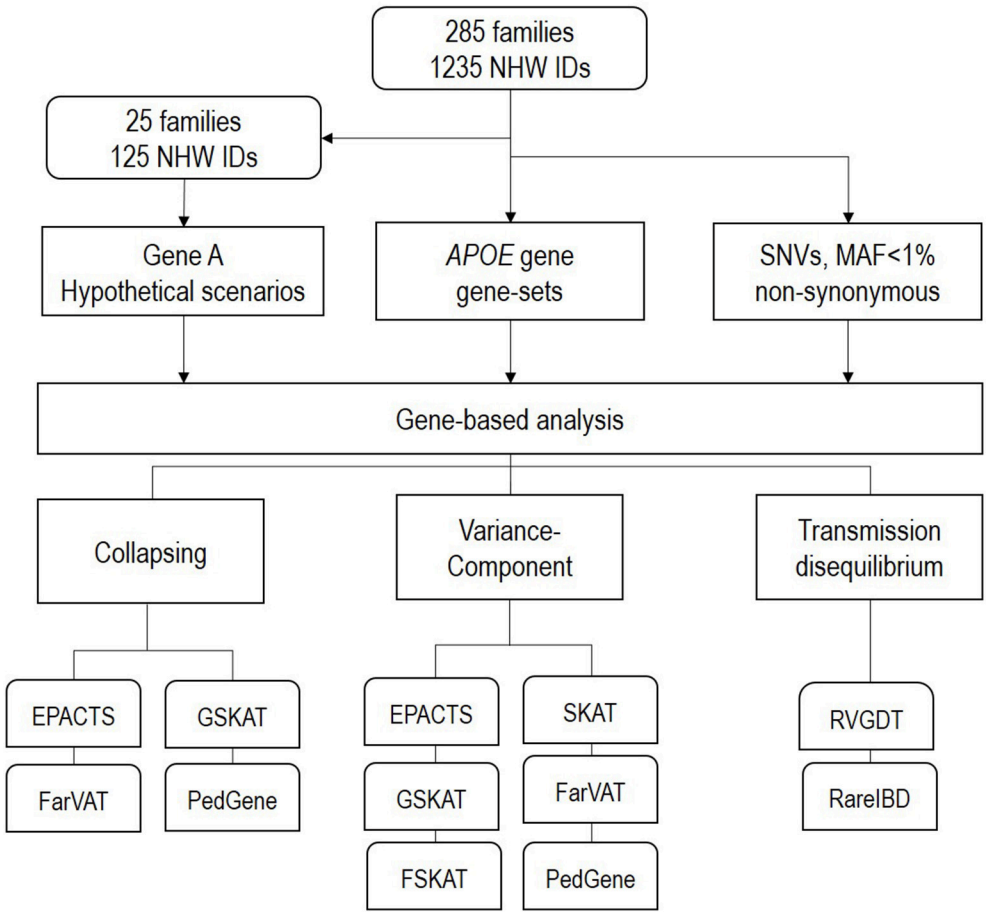

FIGURE 2 | Schematic design of the analysis performed in this study.

TABLE 3 | Relationship of programs and models tested according to their main features and kinship matrix that they use.

	Collapsing			Variance-component		Combined	Transmission-disequilibrium	Kinship		
	Burden	**CMC**	**VT**	**C-ALPHA**	**SKAT**	**SKATO**		**BN**	**IBS**	**Ped**
EPACTS		X	X		X			X		
RVGDT							X			
SKAT-v2					X			X	X	X
GSKAT	X				X					X
FSKAT					X					X
FarVat-Adj	X	X		X		X				
FarVat-BLUP	X	X		X		X				
Pedgne	X				X					
Rarelbd							X			

(scenario 15FC×10FNC), 20 families with carriers and five families who were non-carriers (scenario 20FC×5FNC) and concluded with a scenario in which all 25 families were carriers of one out of the possible five segregating variants in GENE-A (scenario 25FC×0FNC). We tested each of these scenarios with all previously mentioned gene-based methods and software to evaluate their power to associate perfect segregating variants with disease.

Candidate Genes

APOE is the largest genetic risk factor for Alzheimer's disease. The allelic combination of two SNPs, rs429358 (APOE 4;

TABLE 4 | Representation of the segregation pattern of the simulated gene.

	GENE-A				
	SNP1	**SNP2**	**SNP3**	**SNP4**	**SNP5**
Fam1	**1**	0	0	0	0
Fam2	0	**1**	0	0	0
Fam3	0	0	**1**	0	0
Fam4	0	0	0	**1**	0
Fam5	0	0	0	0	**1**
Fam6	**1**	0	0	0	0
Fam7	0	**1**	0	0	0
Fam8	0	0	**1**	0	0
Fam9	0	0	0	**1**	0
Fam10	0	0	0	0	**1**
Fam11	**1**	0	0	0	0
Fam12	0	**1**	0	0	0
Fam13	0	0	**1**	0	0
Fam14	0	0	0	**1**	0
Fam15	0	0	0	0	**1**
Fam16	**1**	0	0	0	0
Fam17	0	**1**	0	0	0
Fam18	0	0	**1**	0	0
Fam19	0	0	0	**1**	0
Fam20	0	0	0	0	**1**
Fam21	**1**	0	0	0	0
Fam22	0	**1**	0	0	0
Fam23	0	0	**1**	0	0
Fam24	0	0	0	**1**	0
Fam25	0	0	0	0	**1**

One (1) means that all cases within the family are carriers of the variant. Zero (0) means that the variant is not present in that family.

19:45411941:T:C), and rs7412 (APOE 2: 19:45412079:C:T), determines one of the three major isoforms of APOE protein, ε2, ε3, or ε4. The dosage of these isoforms determines a person's risk for AD, from having a protective effect in the cases of APOE ε2/ε2 (OR 0.6) or ε2/ε3 (OR 0.6) to different degrees of increased risk according to the number of copies of the ε4 allele (ε2/ε4, OR 2.6; ε3/ε4, OR 3.2; ε4/ε4, OR 14.9) (Farrer et al., 1997). We tested the power of all previously mentioned gene-based methods and software to detect the association of *APOE* gene with disease in our entire dataset ($N = 1{,}235$) under different conditions. We first tested all polymorphic variants (nonsynonymous with MAF $< 1\%$) in the *APOE* gene, next we tested only those variants considered to have a high or moderate effect on the protein including rs429358 and rs7412, then we tested high and moderate effect variants alone, and finally tested rs429358 and rs7412 alone.

Genome-Wide Analyses

We performed gene-based burden analyses on a genome-wide level in our entire dataset (families $n = 285$; samples $N = 1{,}235$) to evaluate the power of each of the previously described methods to detect novel genes significantly associated with disease; only single nucleotide variants (SNVs) with a minor allele frequency equal to or below 1% (MAF $\leq 1\%$), based on the EXAC dataset (Lek et al., 2016), and with a predicted high or moderate effect, according to SnpEff (Cingolani et al., 2012), were included in the analysis. Quantile-Quantile (QQ) plots from gene-based *p*-values were generated with the R package "ggplot2" (Wickham, 2009). We also evaluated the correlations between these methods using Pearson correlation (Pc) and Spearman correlation (Sc) tests of the log of the *p*-values using R v3.4.0 (R Core Team, 2017). Pc evaluates the linear relationship between two continuous variables whereas Sc evaluates the monotonic relationship between two continuous or ordinal variables.

Software Tested

An accompanying supporting file (Supplementary Material) provides a summary of the code employed to run each of the programs described below.

GSKAT

GSKAT (Wang et al., 2013) is among the first R packages developed with the goal of extending burden and kernel-based gene set association tests for population data to related samples with binary phenotypes. To handle the correlated or clustered structure in the family data, GSKAT fits a marginal model with generalized estimated equations (GEE). The basic idea of GEE is to replace the covariance matrix in a generalized linear mix model (GLMM) with a working covariance matrix that reflects the cluster dependencies. Accordingly, GSKAT blends the strengths of kernel machine methods and generalized estimating equations (GEE) to test for the associations between a phenotype and multiple variants in a SNP set. We ran GSKAT correcting for sex and first two PCs.

SKAT

The sequence kernel association test SKAT (Wu et al., 2011) is an R package initially designed for case-control analyses. Later they incorporated the Efficient Mixed-Model Association eXpedited (EMMAX) algorithm (Kang et al., 2010; Zhou and Stephens, 2012) which allows for performing family-based analyses. EMMAX simultaneously corrects for both population stratification and relatedness in an association study by using a linear mixed model with an empirically estimated relatedness matrix to model the correlation between phenotypes of sample subjects. The efficient application of the EMMAX algorithm depends on appropriate estimates of the variance parameters. Relatedness matrices can be calculated based on pedigree structure or estimated from genotype data. For the latter different methods have been proposed. Relatedness can be estimated using those alleles that have descended from a single ancestral allele, i.e., those that are Identical by Descent (IBD), or using the Balding-Nichols (BN) method (Balding and Nichols, 1995) which explicitly models current day populations via their divergence from an ancestral population specified by Wright's F_{st} statistic. We ran SKAT v1.2.1, in R v3.3.3, using the option SKAT_Null_EMMAX correcting for sex and first two PCs and we tested four different kinship matrices: pedigree, IBS, BN and a BN-based kinship matrix (HR) that the EPACTS software constructs (**Table S3**).

FSKAT

FSKAT (Yan et al., 2015), also an R package, is based on a kernel machine regression and can be considered an extension of the sequence kernel association tests (SKAT and famSKAT) for application to family data with dichotomous traits. FSKAT is based on a GLMM framework. Moreover, because it uses all family samples, FSKAT claims to be more powerful than SKAT which uses only unrelated individuals (founders) in the family data. FSKAT constructs a kinship matrix based on pedigree relationships using the R kinship library. We ran FSKAT correcting for sex and first two PCs.

EPACTS

Efficient and Parallelizable Association Container Toolbox (EPACTS) is a stand-alone software that integrates several gene-based statistical tests (CMC, VT, and SKAT) and adapts them to work with complex families by using EMMAX (https://genome.sph.umich.edu/wiki/EPACTS). EPACTS generates a kinship matrix based on the BN algorithm and also annotates the genotypic input file and offers filtering tools (frequency and predicted effect of variants) for easier user-selection of variants that go into gene-based analyses. Nonetheless, we used the same set of variants as in the other tests to run our analysis with EPACTS, correcting for sex and first two PCs.

FarVAT

The Family-based Rare Variant Association Test (FarVAT) (Choi et al., 2014) provides a burden and a variance component test (VT) for extended families and extends these approaches to the SKAT-O statistic. FarVAT assumes that families are ascertained based on the disease status if family members and compares minor allele frequencies between affected and unaffected individuals. FarVAT is implemented with C++ and is computationally efficient. Additionally, if genotype frequencies of affected and unaffected samples are compared to detect genetic associations, it has been shown that the statistical efficiency can be improved by modifying the phenotype; and so FarVAT uses prevalence (Lange and Laird, 2002) or Best Linear Unbalanced Predictor (BLUP) (Thornton and McPeek, 2007) as covariate to modify the genotype.

PedGene

PedGene (Schaid et al., 2013) is an R package that extends burden and kernel statistics to analyze binary traits in family data using large-scale genomic data to calculate pedigree relationships. To derive the kernel association statistic and the burden statistic for data that includes related subjects, they take a retrospective view of sampling with the genotypes considered random.

RVGDT

The Rare Variant Generalized Disequilibrium Test (RVGDT) (He et al., 2017), implemented with Python, differs from the previous methods presented. Instead of using a kernel method to evaluate variants, it uses the generalized disequilibrium test (GDT) which tests genotype differences in all discordant relative pairs to assess associations within a family (Chen et al., 2009). The rare-variant extension of GDT (RVGDT) aggregates a single-variant GDT statistic over a genomic region of interest, which is usually a gene (He et al., 2017). We ran RVGDT correcting for sex and first two PCs.

RareIBD

The developers claim RareIBD (Sul et al., 2016) to be a program without restrictions on family size, type of trait, whether founders are genotyped, or whether unaffected individuals are genotyped. The method is inspired by non-parametric linkage analysis and looks for rare variants with segregation patterns among affected and unaffected individuals that are different from the predicted distributions based on Mendelian inheritance and computes a statistic measuring the difference.

RESULTS

Simulated Dataset

Results from the simulated dataset indicate that RVGDT, rareIBD, and collapsing-based methods (Burden, CMC, and CLP) provided more statistical power than the variance-component methods to detect associations of perfectly segregating variants with disease status (**Table 5**).

In a hypothetical scenario of five families in which each family presented perfect segregation with disease status for a different variant within the same gene (5FC×0NFC), transmission-disequilibrium based methods evaluated this association as significant (even after multiple test correction; e.g., RVGDT $p = 0.004$; p-value after multiple test correction $0.004 \times 9 = 0.036$). RVGDT reached a ceiling p-value of 1×10^{-4}; at 10 families with carriers (FC) plus 15 families of non-carriers (FNC). RVGDT was unable to produce a $p < 9 \times 10^{-4}$, therefore it is not possible to rank or determine the significance of genes that reach this limit. Similarly, RareIBD reports the same p-value for all simulated scenarios, which may be an artifact or a flaw of the program. Collapsing-based methods (Burden, CMC and CLP) started with significant p-values for the 5FC×0NFC scenario, but as we added FNC in the analyses, the associations became less significant. Then as we increased the number of FC of segregating variants, the associations became more significant. In our analyses, most of the variance-component tests could not work with the scenarios containing only five families carrying the segregating variant; most of the tests only provided p-values once 25 families were included in the analyses (5FC×20FNC). After that, as we increased the number of FC of segregating variants, the p-values became smaller. SKAT required 15FC×10FNC to report nominally significant p-values, GSKAT required 20FC×5FNC to report statistically significant p-values, FarVAT-CALPHA did not generate significant p-values unless we used the BLUP correction; FarVAT SKATO reported p-values that were significant at 15FC×10FNC, and at 5FC×20FNC if we used the BLUP correction. P-values from EPACTS-SKAT were not statistically significant after multiple test correction. FSKAT did not deal well with perfectly segregating scenarios; it did not provide p-values for a scenario of only five families all carriers of the segregating variant (5FC×0FNC–FSKAT p-value = NA), and after five families carrying a segregating variant, the program saturated giving no p-value.

Overall, Transmission-disequilibrium tests and collapsing tests were the models that identified the simulated segregating variants as associated with the phenotype; the CMC model provided by FarVAT-BLUP was the one providing most genome-wide significant *p*-values, even in the 5FC×0FNC scenario.

Candidate Genes-APOE

We examined the performance of four gene-sets generated for the *APOE* gene with the 22 family-based gene-based methods in our entire familial cohort. Neither the entire set of polymorphic variants (set "gene" in **Table 6**) nor the set including only rare nonsynonymous variants (set "HM" in **Table 6**) confer risk for these families. The association seems to be driven by the common *APOE* ε2 and ε4 variants, since only when these were included, either alone (set "ε2ε4" in **Table 6**) or in conjunction with the rest of the rare nonsynonymous variants (set "HM-ε2ε4" in **Table 6**) did most of the tests yield a significant *p*-value (after multiple test correction). Only EPACTS-SKAT did not report the *APOE* ε2 and ε4 variants as significantly associated, after multiple test correction, within our dataset (**Table 6**). The most significant association for *APOE* ε2 and ε4 variants was reported by FarVAT-CMC test.

Genome-Wide Analyses

Overall, we examined eight software and over 22 algorithms for genome-wide association analyses in our extended family dataset of 285 families and 1,235 non-hispanic white individuals. We only included in the analyses nonsynonymous SNPs with a MAF $\leq$ 1% and we corrected for sex and first two PCs. All 22 algorithms were run using the same input data. The results for these 22 algorithms are described, grouped per category, in the following sections. First, we compare the correction effect provided by four kinship matrices (**Figure 3A**). Second, we compare the performance of nine variance-component software and algorithms (**Figure 3B**). Third is the comparison of eight collapsing software and algorithms. Fourth, we compare two transmission-disequilibrium tests. We conclude the results section by providing a summary of the pros and cons encountered while running these methods. Overall, most of the results from the gene-based methods tested seemed quite deflated. Only PedGene, FarVAT and Rare-IBD seemed to provide values closer to or above the expected under the null hypothesis. The most efficient in terms of power and *p*-value inflation appears to be FarVAT with BLUP correction.

Kinship Matrices

We tested the correction provided by four kinship matrices using the SKAT method with EMMAX correction implemented in the R package SKATv2. The four kinship matrices tested were pedigree calculation (PED), Identity By State (IBS) estimation, Balding-Nichols (BN) estimation, and the kinship generated by EPACTS (HR) which is also based on the BN algorithm (**Figure 3A**). **Table S3** offers a comparison of these kinships for FAM#1 and FAM#2 of our simulated dataset. For these analyses, we ran the SKAT-EMMAX method in our entire dataset, gene-wide, and calculated a QQ plot and inflation factor (λ) to obtain a general ideal of the behavior of each matrix. Matrices based on the BN algorithm seemed to have a similar performance (SKAT-BN $\lambda = 0.038$, SKAT-HR $\lambda = 0.039$, **Table 7**) though their concordance was lower than expected considering they are based on the same algorithm [Pearson correlation (Pc) = 0.85; Spearman correlation (Sc) = 1]. Although the PED matrix generates a more restrictive correction than the IBS matrix (SKAT-PED $\lambda = 0.36$, SKAT-IBS $\lambda = 0.67$, **Table 7**), these two tests have a similar overall performance as the *p*-values for the different genes were highly correlated ($Pc = 0.97$; $Sc = 0.98$), making the PED matrix a good surrogate for the IBS matrix. Finally, there were clear performance differences between the BN-type matrices (BN and HR) and the IBS-type matrices (IBS and PED), exemplified by the different top candidate genes (*NR1D1* for BN-type matrices and *CHRD* for IBS-type matrices) and by the correlation algorithms (SAKT-IBS vs. SKAT-BN $Pc = 0.8$; $Sc = 0.89$). Overall, we found that the IBS matrix provided the best balance between covariance-correction and overcorrection in our dataset.

Collapsing Tests

The collapsing methods tested from four different software (PedGene, FarVAT, EPACTS and GSKAT) were Burden, CMC, and VT (**Figure 3C**). To compare the different tests we followed a similar approach as above, ran the different software with the same imputed file, and compared the λ.

In our analyses, the burden test by GSKAT presented the most deflated values; though the lambda does not illustrate this (GSKAT-Burden $\lambda = 1.71$, **Table 7**) because of the initial inflation among the low or non-significant genes. EPACTS-CMC ($\lambda = 0.85$) and EPACTS-VT ($\lambda = 0.95$) provided values closer to the expected, and although their QQ-plots appear to follow a similar trend, their correlation is low ($Pc = 0.54$; $Sc = 0.68$) and they reported different top genes. The Burden and CMC methods by FarVAT and FarVAT-BLUP provided *p*-values closest to the expected (FarVAT-Burden $\lambda = 0.98$; FarVAT-CMC $\lambda = 0.99$, FarVAT-BLUP-Burden $\lambda = 1.03$; FarVAT-BLUP-CMC $\lambda = 1.07$). The correlation for the gene *p*-values was higher between results generated by the same method (FarVAT-BLUP-CMC vs. FarVAT-BLUP-Burden $Pc = 0.99$; $Sc = 0.96$; FarVAT-CMC vs. FarVAT-Burden $Pc = 0.98$; $Sc = 0.97$) than between results generated using the same algorithm (FarVAT-BLUP-CMC vs. FarVAT-CMC $Pc = 0.88$; $Sc = 0.8$; FarVAT-BLUP-Burden vs. FarVAT-Burden $Pc = 0.85$; $Sc = 0.77$). PedGene in the burden model was the software that provided the most significant *p*-values; however, these were clearly inflated compared to the predicted *p*-values (Pedgene-Burden $\lambda = 2.99$, **Table 7**) and the results were not correlated with any other Collapsing test (Pc and Sc values < 0.1).

Variance Component Tests

This subset included all the Variance component-based methods available, CLP, CALPHA and SKAT, from six different software: PedGene, FarVAT, FSKAT, EPACTS, SKAT, and GSKAT (**Figure 3C**). GSKAT was the software that reported more deflated values, though the lambda does not illustrate this (GSKAT-SKAT $\lambda = 1.681$, **Table 7**) because of the initial inflation among the low or non-significant genes. GSKAT was followed by SKAT and EPACTS which showed similar λ and performance-values for each gene ($Pc = 0.8$, $Sc =$

TABLE 5 | Gene-based *p*-values for the simulated dataset under different scenarios for the gene-based methods tested in the subset of 25 families.

SET	GSKAT	FSKAT	SKAT	RVGDT	PedGene		Rare IBD	EPACTS*	FarVAT					FarVAT-BLUP				
					SKAT	Burden		SKAT	CMC	CLP	CALPHA	Burden	SKATO	CMC	CLP	CALPHA	Burden	SKATO
5FC×0FNC	0.236	NA	0.141	0.004	0.301	0.003	$<1 \times 10^{-5}$	NA	5.42×10^{-6}	4.66×10^{-6}	NA	NA	NA	3.93×10^{-9}	3.06×10^{-9}	NA	NA	NA
5FC×5FNC	0.235	0.124	0.023	0.002	0.123	7.99×10^{-4}	$<1 \times 10^{-5}$	NA	0.004	0.005	NA	NA	NA	2.10×10^{-5}	4.00×10^{-5}	NA	NA	NA
5FC×10FNC	0.354	0.338	0.112	0.005	0.079	7.99×10^{-4}	$<1 \times 10^{-5}$	NA	0.032	0.036	NA	NA	NA	7.71×10^{-4}	1.01×10^{-3}	NA	NA	NA
5FC×15FNC	0.377	0.359	0.202	0.005	0.095	0.002	$<1 \times 10^{-5}$	NA	0.062	0.061	NA	NA	NA	0.002	2.84×10^{-3}	NA	NA	NA
5FC×20FNC	0.377	0	0.201	0.006	0.114	0.003	$<1 \times 10^{-5}$	0.321	0.073	0.075	0.670	0.075	0.134	0.002	2.40×10^{-3}	0.132	0.002	0.005
10FCA×15FNC	0.083	0	0.028	9×10^{-4}	0.004	2.65×10^{-6}	$<1 \times 10^{-5}$	0.047	0.005	0.008	0.272	0.008	0.017	6.81×10^{-6}	1.33×10^{-5}	0.013	1.33×10^{-5}	3.62×10^{-5}
15FC×10FNC	0.014	0	0.005	9×10^{-4}	0.001	1.77×10^{-9}	$<1 \times 10^{-5}$	0.051	1.72×10^{-6}	6.31×10^{-5}	0.024	6.31×10^{-5}	1.30×10^{-4}	4.26×10^{-11}	3.27×10^{-9}	0.001	3.27×10^{-9}	8.93×10^{-9}
20FC×5FNC	0.002	0	0.002	9×10^{-4}	0.002	1.30×10^{-9}	$<1 \times 10^{-5}$	0.039	1.48×10^{-11}	7.85×10^{-7}	0.024	7.85×10^{-7}	1.14×10^{-6}	6.12×10^{-18}	2.12×10^{-12}	6.32×10^{-4}	2.12×10^{-12}	2.54×10^{-10}
25FC×0FNC	3×10^{-4}	0	0.001	9×10^{-4}	0.001	1.42×10^{-10}	$<1 \times 10^{-5}$	0.033	1.55×10^{-19}	4.44×10^{-8}	0.025	4.44×10^{-8}	7.06×10^{-8}	4.59×10^{-29}	4.58×10^{-15}	5.10×10^{-4}	4.58×10^{-15}	2.54×10^{-10}

Simulated scenarios: 5FC, five families carrier of variants within the hypothetical gene; 5FC×5FNC, five families carrier of variants within the hypothetical gene and five families non-carrier of variants within the hypothetical gene; 5FC×10FNC, five families carrier of variants within the hypothetical gene and 10 families non-carrier of variants within the hypothetical gene; 5FC×15FNC, five families carrier of variants within the hypothetical gene and 15 families non-carrier of variants within the hypothetical gene; 5FC×20FNC, five families carrier of variants within the hypothetical gene and 20 families non-carrier of variants within the hypothetical gene; 10FC×15FNC, 10 families carrier of variants within the hypothetical gene and 15 families non-carrier of variants within the hypothetical gene; 15FC×10FNC, 15 families carrier of variants within the hypothetical gene and 10 families non-carrier of variants within the hypothetical gene; 20FC×5FNC, 20 families carrier of variants within the hypothetical gene and five families non-carrier of variants within the hypothetical gene; 25FC, 25 families carrier of variants within the hypothetical gene.

**we tested SKAT, CMC, and VT on EPACTS, but CMC and VT reported all NA values so data is not shown.*

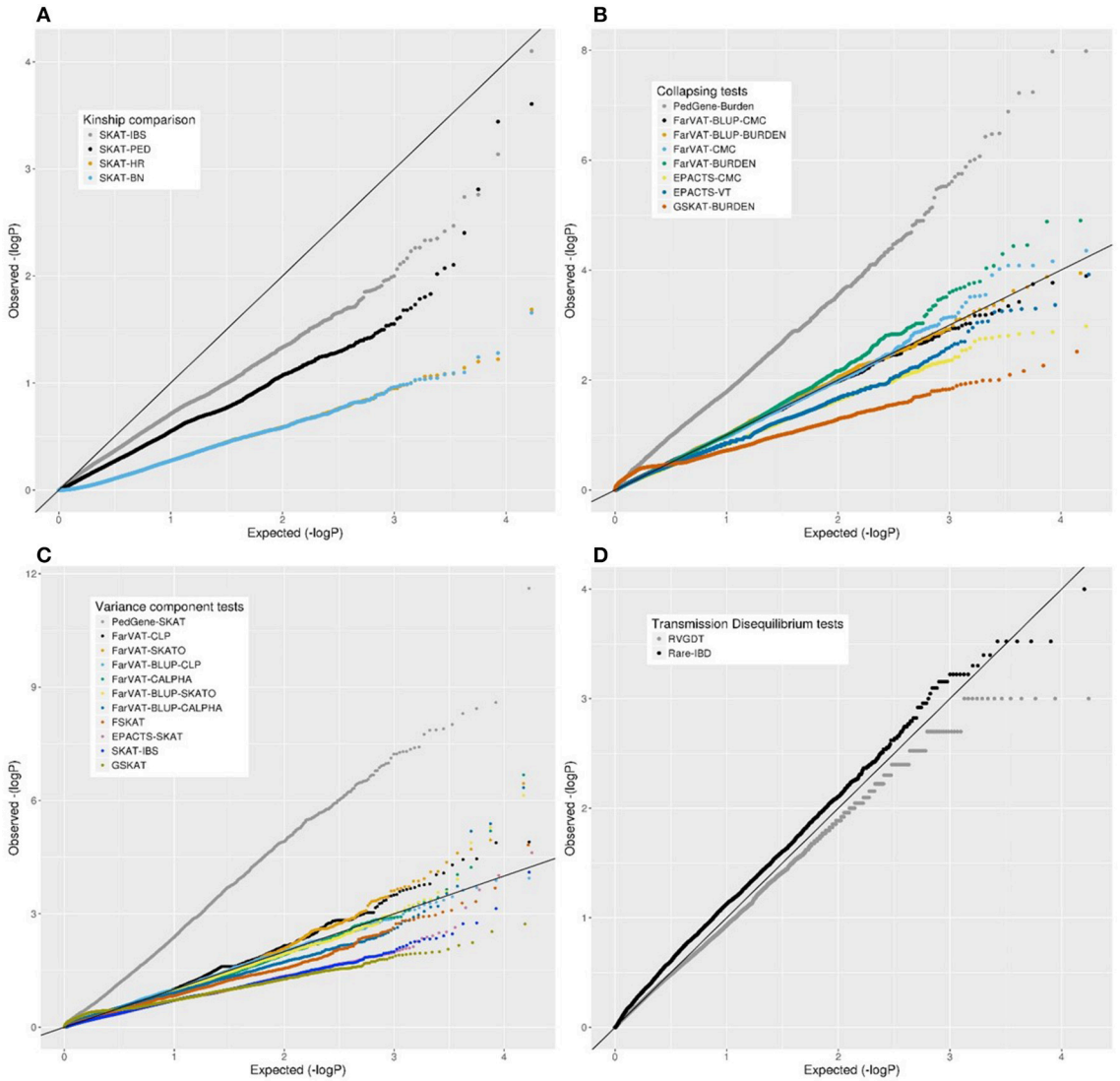

FIGURE 3 | Quantile-quantile (QQ) plots from different family-based gene-based methods for all nonsynonymous variants with a MAF <1% in our family-based dataset. **(A)** Comparison of SKAT test using different kinship matrices: pedigree calculation (PED), Identity By Similarity (IBS) estimation, Balding-Nichols (BN) estimation, and the kinship generated by EPACTS (HR). **(C)** Comparison of different collapsing tests: GSKAT, EPACTS, FarVAT, and PedGene. **(B)** Comparison of different variance-component gene-based methods: GSKAT, FSKAT, SKAT, EPACTS, FarVAT, and PedGene. **(D)** Comparison of transmission disequilibrium tests: RVGDT and RareIBD.

0.8, **Figure 4**). The CLP, CALPHA, and SKATO methods by FarVAT and FarVAT-BLUP provided *p*-values closest to the expected (FarVAT-CLP $\lambda = 1.00$; FarVAT-CALPHA $\lambda = 1.15$; FarVAT-SKATO $\lambda = 1.02$, FarVAT-BLUP-CLP $\lambda = 1.11$; FarVAT-BLUP-CALPHA $\lambda = 1.26$; FarVAT-BLUP-SKATO $\lambda = 1.10$). FarVAT-CALPHA, FarVAT-SKATO, FarVAT-BLUP-CALPHA and FarVAT-BLUP-SKATO reported the same top candidate gene (*CHRD*) (**Table 7**), though the overall *p*-value correlation was lower than expected considering they are based on the same algorithm (FarVAT-SKATO vs. FarVAT-BLUP-SKATO $Pc = 0.6$, $Sc = 0.7$; FarVAT-CALPHA vs. FarVAT-BLUP-CALPHA $Pc = 0.82$ $Sc = 0.82$, **Figure 4**). On the other hand, despite the fact that FarVAT-CLP and FarVAT-BLUP-CLP had higher correlation ($Pc = 0.85$, $Sc = 0.77$), these two tests reported different top genes (FarVAT-CLP top gene is *MAS1L*, and FarVAT-BLIP-CLP top gene is *NLRP9*). PedGene in the SKAT model was the software that provided the most significant *p*-values, but these were clearly inflated (Pedgene-SKAT $\lambda = 3.53$, **Table 7**) and its correlation with other variance component tests was low to null (Pc and Sc values < 0.2).

Transmission Disequilibrium Tests

We tested two transmission disequilibrium tests, RVGDT and Rare-IBD, which were designed to account for large extended families of arbitrary structure (**Figure 3D**). Of these two, RVGDT was the test that more closely approached the expected under the

TABLE 6 | Gene-based *p*-values for the *APOE* gene under different gene-set scenarios for the gene-based methods tested in the entire dataset ($N = 1235$, 285 families).

APOE	N	GSKAT	FSKAT	SKAT	RVGDT	PedGene		Rare IBD	EPACTS*	FarVAT				FarVAT-BLUP				
						SKAT	Burden		SKAT	CMC	CLP	CALPHA	Burden	SKATO	CMC	CLP	CALPHA	Burden	SKATO
gene	19	0.035	0.037	0.061	0.164	**0.008**	0.515	0.712	0.205	0.053	0.379	**0.003**	0.379	**0.005**	0.036	0.311	0.017	0.311	0.034
HM-ε2ε4	4	**0.003**	**0.002**	**0.001**	**0.005**	0.412	0.414	0.359	0.020	$\mathbf{7.87 \times 10^{-15}}$	0.420	$\mathbf{4.99 \times 10^{-4}}$	0.420	**0.001**	$\mathbf{3.73 \times 10^{-14}}$	0.275	$\mathbf{3.99 \times 10^{-4}}$	0.275	$\mathbf{6.99 \times 10^{-4}}$
HM	2	0.067	0.089	0.048	0.237	0.177	0.177	0.741	0.022	0.028	0.052	0.014	0.052	0.018	0.053	0.090	0.024	0.090	0.031
ε2ε4	2	**0.005**	**0.002**	**0.003**	**0.004**	0.849	0.855	**0.002**	0.024	$\mathbf{7.87 \times 10^{-15}}$	**0.002**	**0.002**	**0.002**	**0.003**	$\mathbf{3.73 \times 10^{-14}}$	**0.002**	**0.001**	**0.001**	**0.001**

In the analysis, only nonsynonymous variants (only SNVs) with a MAF<0.01, and the APOE ε2 and ε4, were considered and we adjusted by sex and PCAs. Highlighted in bold, significant p-values after multiple test correction. gene, set of 19 polymorphic variants within APOE gene, including APOE ε2 and ε4 variants; HM-ε2ε4, set of variants considered HIGH or MODERATE including APOE ε2 and ε4 variants; HM, set of variants considered HIGH or MODERATE without APOE ε2 and ε4 variants; ε2ε4, APOE ε2 and ε4 variants alone. N, number of variants that went into analysis.

**We tested SKAT, CMC, and VT on EPACTS, but CMC and VT reported all NA values so data is not shown.*

null ($\lambda = 0.99$), whereas Rare-IBD provided slightly inflated *p*-values ($\lambda = 1.450$, **Table 7**). The correlation between these two methods was very low (Pearson correlation = 0.23, Spearman correlation = 0.17). A common issue with both methods was that we observed some stratification toward more significant *p*-values which made it difficult to determine a top significant gene.

Pros and Cons of the Different Gene-Based Methods

Among all the methods tested, EPACTS and FarVAT are the most user-friendly, time-efficient and versatile software. EPACTS is an all-in-one package that annotates the input file, generates the kinship matrix and performs gene-based analysis under different conditions (minor allele frequency and predicted functionality of the variant) with only tag specification. In addition, the program can be run on a genome-wide basis or at a smaller scale given genes or regions specified by the user. FarVAT can generate the kinship matrix by either using the pedigree relationships or using the genetic relationship among individuals. It does not annotate the input file and requires that the user provide their own set of genes and variants per gene to analyze; it allows the user to choose between BLUP or prevalence to estimate and incorporate random effects on the phenotype. FarVAT has initial conditioning that only takes founder-based MAF, so when a genetic variant only has minor alleles in non-founders (offspring) these numbers will not be counted. This is a big limitation with respect to the other programs that take into account all variants regardless of their presence in founders or not. Since we only had genetic data for siblings for many of our families, so no genetic data for founders, we ran FarVAT with the "–freq all" option so that all variants would be included regardless if they were present in founders or not.

FSKAT, GSKAT, and SKAT require some R knowledge from the user, and are less flexible. For FSKAT and GSKAT the user has to provide a genotype, a phenotype, and a gene-set file. For SKAT the user has to additionally provide the kinship matrix. Because these programs were designed to run on a per gene basis, these take longer computational time to be run on a genome-wide level than EPACTS or FarVAT, even if the user parallelizes computation. PedGene is also an R package that requires a genotype, a phenotype file with complete pedigree information (to generate the kinship matrix), and a gene-set file. PedGene provides phenotype adjustment by logistic regression on the trait of interest, but it does not allow for extra covariates, which prohibits correction by multiple PCs or other variables. RVGDT is a Python based program, quite user-friendly since it is operated with simple command-line but is limited in its options. Similar to FSKAT, GSKAT, and SKAT, it is designed to be run on a per-gene basis for which loops and parallelization have to be set up for genome-wide testing. The same applies to RareIBD which requires a genotype, a phenotype, and a Kinship coefficient file for each gene that the user wants to test. For each gene the program first computes statistics for each founder within each family and then calculates the gene-based *p*-value. The first step of this process can easily take between 3 and 5 min for families with <100 individuals; hence, the overall time for one gene is directly dependent on the number of families and the time required for a genome-wide analysis is proportional to the number of genes

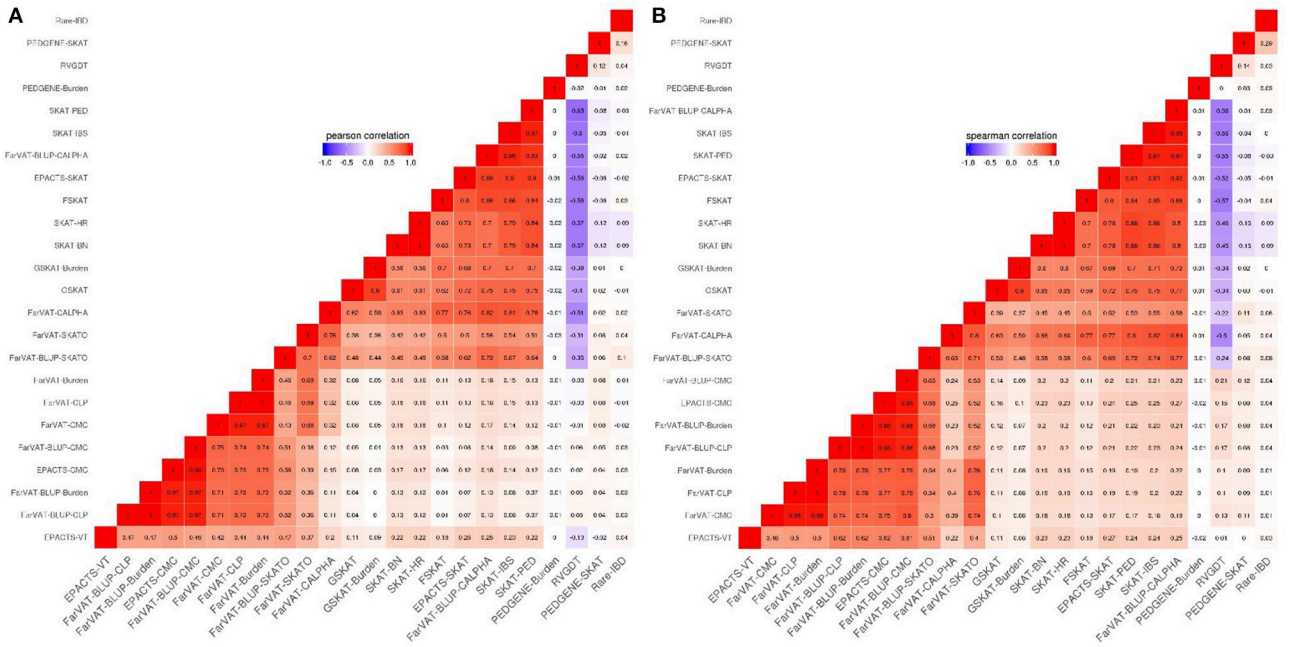

FIGURE 4 | Correlation plots from different family-based gene-based methods for genes with a $p \leq 0.005$. **(A)** Pearson correlation correlates genes according to their *p*-values. **(B)** Spearman correlation correlates genes according to their rankings.

being tested. Although it is possible to parallelize the jobs using a high-performance cluster (if available) this program is the slowest of all tested.

One of the major drawbacks we found is that some of these programs do not accept missing data (FSKAT or RareIBD) or will not generate a *p*-value if the gene set contains only one variant (GSKAT, PedGene or FarVAT). FSKAT does not accept missing data, and although it calculated *p*-values for genes that only have one informative SNP (one-SNP-gene), there were at least 75 (3.26%) of 2,154 one-SNP-genes for which the returned *p*-value was "2." GSKAT did not provide *p*-values for more than 1,875 one-SNP-genes. PedGene also had trouble generating *p*-values for 44 one-SNP-genes out of a total of 1,916 singletons. FarVAT did not generate *p*-values for the one-SNP-genes using the Burden and SKATO models but it did generate *p*-values using the CMC and CLP models for the same 1,875 one-SNP-genes.

Candidate Genes for FASe Project

Our results indicate that transmission disequilibrium tests identify genes that have a Mendelian behavior, whereas collapsing and variance-component tests identify genes that confer risk for disease. Therefore, we decided to combine and compare results from all approaches to identify the genes with most consistent results (**Table 8**).

PedGene provided the most significant *p*-values for *NTN5* (Pedgene-Burden $p = 5.80 \times 10^{-8}$; PedGene-SKAT $p = 1.26 \times 10^{-8}$) and *ANKRD42* (PedGene-Burden $p = 3.62 \times 10^{-7}$; PedGene-SKAT $p = 1.16 \times 10^{-7}$). However, the inflated *p*-values observed and low correlation with any of the other software tested using the same algorithms makes us suspicious of the validity of these results.

CHRD was the gene with the third most significant *p*-value. *CHRD* had a $p \leq 5 \times 10^{-7}$ in three different models (FarVAT-CALPHA, FarVAT-SKATO, and FarVAT-BLUP-CALPHA). Additionally, as we lowered the considered *p*-value threshold, we found that more tests identified *CHRD* as a potential candidate gene associated with AD. When we lowered the threshold to suggestive genome-wide *p*-value ($p \leq 5 \times 10^{-4}$) we found that seven different models identified *CHRD* as significantly associated with AD. Following the same method we found that *CLCN2*, *MAS1L*, and *PTK2B* had $p \leq 5 \times 10^{-05}$ in at least three tests, and if we lowered the threshold to $\leq 5 \times 10^{-4}$ *p*-value, these genes were identified as significant by at least three additional tests.

Among genes with a $p \leq 5 \times 10^{-04}$; *CPAMD8* was identified by at least nine gene-based methods (FarVAT, FarVAT-BLUP, and PedGene). The exact *p*-value for *CPAMD8* could not be estimated by RVGDT as it reported a *p*-value of 9×10^{-04}, which is the most significant *p*-value reported by this test. Therefore, we cannot conclude that *CPAMD8* presented a *p*-value $\leq 5 \times 10^{-04}$ by RVGDT. *CHRD*, *CLCN2, MAS1L*, *PTK2B*, and *CPAMD8*, *NLRP9*, and *HDLBP* were also potential novel candidate genes for familial LOAD as they had $p \leq 5 \times 10^{-04}$ using at least five or more tests (**Table 8**).

Since these were identified by multiple gene-based methods, we wanted to determine whether any of these seven candidate genes are involved in known AD pathways. Common variants in *PTK2B* have been associated with AD risk at a genome-wide level (Lambert et al., 2013). Our results indicate there are additional

TABLE 7 | Top results for all gene-based methods tested.

Software	TEST	Top gene	Top *p*-value	Lambda
PedGene	SKAT	*KANSL1L*	2.42×10^{-12}	3.533
PedGene	Burden	*TTN*	1.04×10^{-8}	2.997
GSKAT	Burden	*PCSK6*	3.04×10^{-3}	1.704
GSKAT	SKAT	*NR1D1*	1.90×10^{-3}	1.681
Rare-IBD	TDT	*SNTB2*	1.00×10^{-4}	1.450
FarVAT-BLUP	CALPHA	*CHRD*	4.60×10^{-07}	1.259
FarVAT	CALPHA	*CHRD*	2.09×10^{-07}	1.152
FarVAT-BLUP	CLP	*NLRP9*	1.14×10^{-4}	1.112
FarVAT-BLUP	SKATO	*CHRD*	7.37×10^{-7}	1.101
FarVAT-BLUP	CMC	*IGHV1-69*	1.28×10^{-4}	1.066
FarVAT-BLUP	Burden	*NLRP9*	1.14×10^{-4}	1.031
FarVAT	SKATO	*CHRD*	3.54×10^{-7}	1.016
FarVAT	CLP	*MAS1L*	1.25×10^{-5}	1.000
RVGDT	TDT	*RTN3*	9.99×10^{-4}	0.995
FarVAT	CMC	*HSD3B1*	4.40×10^{-5}	0.993
FarVAT	Burden	*MAS1L*	1.25×10^{-5}	0.985
EPACTS	VT	*PPAN-P2RY11*	1.20×10^{-4}	0.954
FSKAT	SKAT	*CHRD*	2.00×10^{-5}	0.938
EPACTS	CMC	*BTN2A2*	1.05×10^{-3}	0.849
SKAT	IBS	*CHRD*	7.94×10^{-5}	0.668
EPACTS	SKAT	*CHRD*	2.42×10^{-5}	0.635
SKAT	PED	*CHRD*	2.47×10^{-4}	0.360
SKAT	HR	*NR1D1*	2.06×10^{-2}	0.039
SKAT	BN	*NR1D1*	2.21×10^{-2}	0.038

Top gene, p-value and lambda for each test is given, ordered by lambda value.

low-frequency and rare nonsynonymous variants in *PTK2B* that are associated with AD risk in late-onset families.

We used the GeneMANIA (http://pages.genemania.org/) algorithm on the seven candidate genes (*CHRD*, *MAS1L*, *PTK2B, CPAMD8*, *NLRP9*, *CLCN2*, and *HDLBP*) and known AD-related genes (*APP, PSEN1, PSEN2, APOE, TREM2, PLD3,* and *ADAM10*) which are involved in some pathways important in AD (APP-metabolism and immune response). GeneMANIA looks for relationships among a list of given genes by searching within multiple publicly available biological datasets. These datasets include protein-protein, protein-DNA and genetic interactions, pathways, reactions, gene and protein expression data, protein domains and phenotypic screening profiles. We found that our candidate genes have genetic interactions and co-localization with known AD genes. *CHRD* and *PTK2B* are involved in "regulation of cell adhesion" like *ADAM10*; *PTK2B* is involved in "regulation of neurogenesis" like *APOE* and "perinuclear region of cytoplasm" like *APP*, *PSEN1* and *PSEN2*. Finally, *CLCN2* and *PTK2B* are connected through "regulation of ion transport" (**Figure 5**).

DISCUSSION

The missing heritability in AD, and in many complex diseases, may be found in very rare variants for which discovery will require either large datasets (e.g., the ADSP Discovery Phase which has over 10,000 sequenced individuals) or datasets enriched for rare variants (such as families with history of AD). In this study, we present the most comprehensive performance analyses of multiple gene-based methods using 285 families with AD. Some of the current methods available are underpowered or too restrictive to detect genes significantly associated with this disease (**Figure 4**). Results from our simulated data (**Table 5**) show that only certain highly-restricted scenarios provide gene-wide significant *p*-values in family-based analyses; whereas similar scenarios in a case-control study would result in gene-wide *p*-values. To circumvent this power issue, we relied on the combination of multiple evidence toward the same gene.

One key aspect to adapt gene-based analyses to a family-based context is to account for population stratification and hidden relatedness that may appear due to the inherent nature of family datasets. To take into account this issue, gene-based algorithms must incorporate kinship matrices to model the relationships among samples. Therefore, an appropriate estimate of the kinship matrix is of utmost importance. In this work we show how different relationship matrices influence results. We tested the three most common types of kinship matrix, pedigree reconstruction (PED), identity by state (IBS), and Balding-Nichols (BN). We show that for a situation of complex incomplete families, correction using PED or BN matrices will lead to an overcorrection of the relationships decreasing the power of these tests (**Table 7**, **Figure 4A**).

In order to choose the best gene-based algorithm for analysis, it is important to take into account the nature (impact and directionality) of the variants that are being included in the test. Collapsing tests are powerful when a large proportion of variants are causal and the effects are in the same direction. Variance-component tests are supposed to be more powerful than collapsing tests because they allow for admixture of risk and protective variants within the region being tested (Ionita-Laza et al., 2013). It is not practical to account for the nature of the variants included in each gene-set, and the true disease model is unknown and variable; hence, omnibus or combined tests such as SKAT-O would be desirable for genome-wide studies (Lee et al., 2012). However, most family-based methods do not incorporate the SKAT-O algorithm, except FarVAT. Therefore, the best approach to perform genome-wide rare variant discovery is to combine different algorithms and look for common signatures across the tests performed. Nonetheless, we are aware that running all available tests is a time-consuming task that requires additional expertise and resources. In our analyses FarVAT, with the BLUP adjustment, provide the best results in terms of significant *p*-values and minor inflation, for genome-wide gene-based analysis; it is a fast software that provides results from multiple tests at the same time. The R version of SKAT or EPACTS would be alternatively valid choices, taking into account that these overcorrect and the *p*-value threshold should be lowered.

In this study, we identified *CHRD* as a candidate gene with a genome-wide significant *p*-value (5×10^{-07}) reported by three tests, and another six genes that had a suggestive

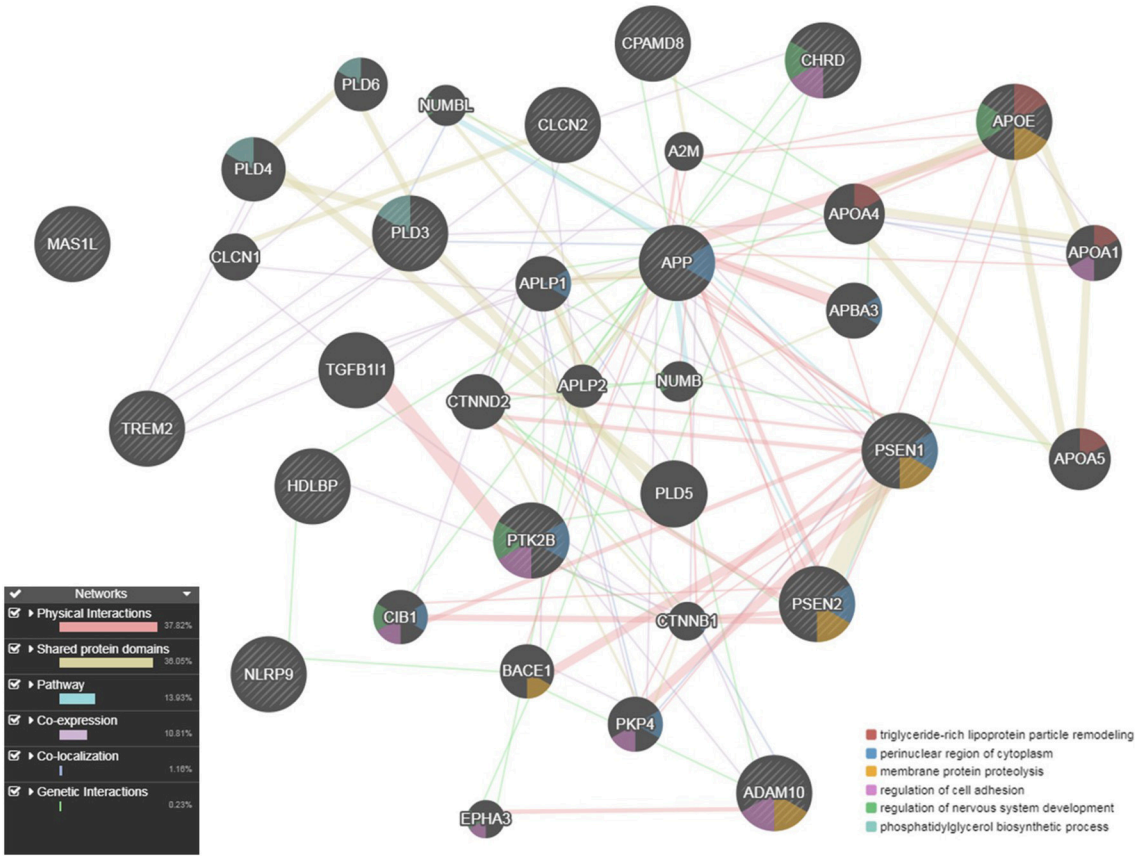

FIGURE 5 | Gene network for the seven candidate genes (*CHRD, CLCN2, CPAMD8, HDLBP, MAS1L, NLRP9*, and *PTK2B*) with multiple evidence of a $p \leq 5 \times 10^{-04}$, anchored with known AD genes (*APP, PSEN1, PSEN2, APOE, TREM2, ADAM10,* and *PLD3*), as described by GeneMANIA.

genome-wide $p < 5 \times 10^{-04}$ in at least five, and up to nine, of the different test performed: *CLCN2, CPAMD8, HDLBP, MAS1L, NLRP9*, and *PTK2B*. Additionally, these genes seem to have direct and indirect interactions (genetic interaction, co-localization or shared function) with known AD genes (*APP, PSEN1, PSEN2, APOE, TREM2, PLD3*, and *ADAM10*).

CHRD, chordin, is a highly-conserved developmental protein which inhibits the ventralizing activity of bone morphogenetic proteins, is active during gastrulation, expressed in fetal and adult liver and cerebellum, and is associated with Cornelia de Lange syndrome (Smith et al., 1999). *CLCN2,* chloride voltage-gated channel 2, has several functions including the regulation of cell volume: membrane potential stabilization, signal transduction and transepithelial transport. It has been associated with different epilepsy modes (Saint-Martin et al., 2009; Cukier et al., 2014) and leukoencephalopathy (Gaitán-Peñas et al., 2017). *CHRD* and *CLCN2* show co-expression which could be due to their close proximity, both belong to a gene cluster at 3q27. Interestingly, *CLCN2* shows co-expression with *TREM2* which, other than being an AD risk gene, is known to cause leukoencephalopathy in PLOSL (polycystic lipomembranous osteodysplasia with sclerosing leukoencephalopathy), also known as Nasu-Hakola disease.

PTK2B, Protein Tyrosine Kinase 2 Beta, was described as an AD risk locus in the largest GWAS meta-analysis conducted to date (Lambert et al., 2013), and later corroborated by others (Beecham et al., 2014; Wang et al., 2015). The protein encoded by *PTK2B* is a member of the focal adhesion kinase (FAK) family that can be activated by changes in intracellular calcium levels, which are disrupted in AD brains. Its activation regulates neuronal activity such as mitogen-activated protein kinase (MAPK) signaling (Rosenthal and Kamboh, 2014). *PTK2B* could also be involved in hippocampal synaptic function (Lambert et al., 2013). Although there is no co-expression or genetic interaction between *CLCN2* and *PTK2B*, both are involved in regulation of ion transport. Additionally, *PTK2B* is involved in regulation of lipidic metabolic processes like *APOE*, a cholesterol-related gene. Although no association has yet been reported between *APOE* and *HDLBP*, the High-Density Lipoprotein Binding Protein, the latter plays a role in cell sterol metabolism, protecting cells from over-accumulation of cholesterol, which has been reported as risk factor for atherosclerotic vascular diseases.

CPAMD8, C3 and PZP Like, Alpha-2-Macroglobulin Domain Containing 8, has been previously associated with neurological conditions other than AD. Common variants in *CPAMD8* were found among top markers associated with multiple sclerosis

TABLE 8 | Most frequent genes, within *p*-value threshold category, across the different gene-based family-based methods tested*.

P-value threshold	gene	No.	EPACTS			FSKAT	GSKAT		RVGDT	SKAT	FarVAT				FarVAT-BLUP					Rare-IBD
			CMC	VT	SKAT		SKAT	Burden		IBS	CMC	CLP	Burden	CALPHA	SKATO	CMC	CLP	Burden	CALPHA	SKATO	
$\leq 5 \times 10^{-7}$	*CHRD*	3	0.007	0.031	2.42×10^{-5}	1.50×10^{-5}	0.013	0.013	0.990	7.94×10^{-5}	0.007	0.007	0.007	**2.09×10^{-7}**	**3.54×10^{-7}**	0.004	0.004	0.004	**4.06×10^{-7}**	7.37×10^{-7}	0.071
$\leq 5 \times 10^{-6}$	*CHRD*	4	0.007	0.031	0.000	0.000	0.013	0.013	0.980	0.000	0.007	0.007	0.007	**2.09×10^{-7}**	**3.54×10^{-7}**	0.004	0.004	0.004	**4.06×10^{-7}**	**7.37×10^{-7}**	0.071
$\leq 5 \times 10^{-5}$	*CHRD*	5	0.007	0.031	**2.42×10^{-5}**	**1.50×10^{-5}**	0.013	0.013	0.990	0.000	0.007	0.007	0.007	**2.09×10^{-7}**	**3.54×10^{-7}**	0.004	0.004	0.004	**4.06×10^{-7}**	**7.37×10^{-7}**	0.071
	CLCN2	4	0.018	0.043	2.33×10^{-4}	2.07×10^{-4}	0.002	0.020	1.000	7.30×10^{-4}	0.006	0.005	0.005	**6.46×10^{-6}**	**1.12×10^{-5}**	0.011	0.009	0.009	**6.51×10^{-6}**	**1.32×10^{-5}**	0.299
	MAS1L	3	0.002	0.003	0.057	0.019	0.187	0.187	0.998	0.042	4.65×10^{-4}	**1.25×10^{-5}**	**1.25×10^{-5}**	4.27×10^{-4}	**1.96×10^{-5}**	0.001	1.32×10^{-4}	1.32×10^{-4}	0.015	2.73×10^{-4}	0.685
	PTK2B	3	0.001	0.009	0.331	0.205	0.090	0.090	1.000	0.193	1.23×10^{-4}	**1.31×10^{-5}**	**1.31×10^{-5}**	0.060	**2.46×10^{-5}**	0.001	2.39×10^{-4}	2.39×10^{-4}	0.113	4.93×10^{-4}	0.443
$\leq 5 \times 10^{-4}$	*CPAMD8*	8	0.002	0.003	0.652	0.178	0.155	0.191	9.99×10^{-4}	0.572	**6.91×10^{-5}**	**2.02×10^{-4}**	**2.02×10^{-4}**	0.309	**4.22×10^{-4}**	**1.69×10^{-4}**	**2.03×10^{-4}**	**2.03×10^{-4}**	0.268	**4.23×10^{-4}**	6.00×10^{-4}
	NLRP9	8	0.001	0.013	0.020	0.013	0.029	0.029	0.998	0.019	**2.81×10^{-4}**	**2.40×10^{-4}**	**2.40×10^{-4}**	0.002	**3.78×10^{-4}**	**4.50×10^{-4}**	**1.14×10^{-4}**	**1.14×10^{-4}**	0.003	**2.59×10^{-4}**	0.157
	MAS1L	8	0.002	0.003	0.057	0.019	0.187	0.187	0.998	0.042	**4.65×10^{-4}**	**1.25×10^{-5}**	**1.25×10^{-5}**	**4.27×10^{-4}**	**1.96×10^{-5}**	0.001	**1.32×10^{-4}**	**1.32×10^{-4}**	0.015	**2.73×10^{-4}**	0.685
	CHRD	7	0.007	0.031	**2.42×10^{-5}**	**1.50×10^{-5}**	0.013	0.013	0.990	**7.94×10^{-5}**	0.007	0.007	0.007	**2.09×10^{-7}**	**3.54×10^{-7}**	0.004	0.004	0.004	**4.60×10^{-7}**	**7.37×10^{-7}**	0.071
	PTK2B	7	0.001	0.009	0.331	0.205	0.090	0.090	1.000	0.193	**1.23×10^{-4}**	**1.31×10^{-5}**	**1.31×10^{-5}**	0.060	**2.46×10^{-5}**	0.001	**2.39×10^{-4}**	**2.39×10^{-4}**	0.113	**4.93×10^{-4}**	0.443
	CLCN2	6	0.018	0.043	**2.33×10^{-4}**	**2.07×10^{-4}**	0.020	0.020	1.000	7.30×10^{-4}	0.006	0.005	0.005	**6.46×10^{-6}**	**1.12×10^{-5}**	0.011	0.009	0.009	**6.51×10^{-6}**	**1.32×10^{-5}**	0.299
	HDLBP	5	0.002	0.024	0.009	0.001	0.031	0.032	0.996	0.002	0.021	0.028	0.028	0.068	0.046	**1.79×10^{-4}**	**4.92×10^{-4}**	**4.92×10^{-4}**	**2.89×10^{-4}**	**1.22×10^{-4}**	0.428

Highlighted in bold the tests with significant p-value according to threshold category.

**PedGene results have not been included given the inflated results of this test and the low correlation with the other gene-based methods.*

(Baranzini et al., 2009). Missense and frameshift variants in *CPAMD8* were identified in three families affected with Anterior Segment Dysgenesis (Cheong et al., 2016). According to the UKBiobank PheWeb (http://pheweb.sph.umich.edu:5000/), *CPAMD8* has a 2.9×10^{-9} *p*-value for its association with AD. We did not find any shared pathway between *CPAMD8* and known AD genes in the GeneMANIA network, even though it seems to have a genetic interaction with *APP* (Lin et al., 2010). In our study *CPAMD8* was identified as a candidate gene (with $p < 1 \times 10^{-4}$) for AD by at least nine gene-based methods from different software, and we found that several variants within this gene had varying degrees of segregation in more than twenty families. Variant p.(Ser1103Ala) segregates with disease status in two families with two and three carriers respectively, and is present in another two families. Variant p.(His465Arg) segregates with disease status in five families with two or three carriers per family and is present in another 11 families. Variant p.(Arg1380Cys) is private to a family with three carriers, p.(Ala1492Pro) is private to a family with five carriers, and p.(Val521Met) is private to a family with three carriers.

MAS1L, MAS1 Proto-Oncogene Like, is a G Protein-Coupled Receptor. Members of this family of membrane proteins are activated by a wide spectrum of ligands and modulate the activity of different signaling pathways in a ligand-specific manner. Aly et al. (2008) described polymorphisms in the region of the UBD/MAS1L genes that are associated with type-1 diabetes.

The immune system and the integrity of the blood-brain barrier are key factors for Alzheimer disease. *NLRP9*, NLR Family Pyrin Domain Containing 9, has been involved in inflammation response. Nyúl-Tóth et al. (2017) found *NLRP9* expressed in cerebral endothelial cells and, at much lower levels, in brain pericytes; and another member of the NLP family (*NLRP1*) has been associated with AD (Pontillo et al., 2012).

We have reviewed more than 22 algorithms from eight different software available for gene-based analyses in complex families. After a thorough examination of the performance of these tests under different scenarios, we present a methodology to identify genes associated with the studied phenotype. We have applied this methodology to 285 European-American families affected with late onset Alzheimer disease (LOAD) and we identified six candidate genes with suggestive or genome-wide significant *p*-values across different software and algorithms. Based on the consistency of our results, we are confident that some of these genes may play a role in AD pathology and therefore are of interest to follow up in replication and functional studies.

AUTHOR CONTRIBUTIONS

MF performed processing and quality control of data, implementation of software and statistical analysis; participated in study design, interpretation of results and wrote the manuscript. JB contributed to data collection, processing and quality control. JB, JD-A, LI, YD, and OH contributed on software implementation and interpretation or results. JN, JM, and AG contributed to study participant recruitment and sample

collection. CC collected data, participated in study design, interpretation of results and revision of manuscript. All the authors read and provided input to the manuscript.

FUNDING

This work was supported by grants from the National Institutes of Health (R01-AG044546, P01-AG003991, and RF1-AG053303), the Alzheimer Association (NIRG-11-200110, BAND-14-338165 and BFG-15-362540) and the JPB Foundation. The recruitment and clinical characterization of research participants at Washington University were supported by NIH P50-AG05681, P01-AG03991, and P01-AG026276. Samples from the National Cell Repository for Alzheimer's Disease (NCRAD), which receives government support under a cooperative agreement grant (U24-AG21886) awarded by the National Institute on Aging (NIA), were used in this study. NIALOAD samples were collected under a cooperative agreement grant (U24-AG026395) awarded by the National Institute on Aging.

We thank the Genome Technology Access Center in the Department of Genetics at Washington University School of Medicine for help with genomic analysis. The Center is partially supported by NCI Cancer Center Support Grant #P30 CA91842 to the Siteman Cancer Center and by ICTS/CTSA Grant# UL1TR000448 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research. This work was supported by access to equipment made possible by the Hope Center for Neurological Disorders and the Departments of Neurology and Psychiatry at Washington University School of Medicine.

ACKNOWLEDGMENTS

We thank contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible. Members of the National Institute on Aging Late-Onset Alzheimer Disease/National Cell Repository for Alzheimer Disease (NIA-LOAD NCRAD) Family Study Group include the following: Richard Mayeux, M.D., M.Sc., Martin Farlow, M.D., Tatiana Foroud, Ph.D., Kelley Faber, M.S., Bradley F. Boeve, M.D., Neill R. Graff-Radford, M.D., David A. Bennett, M.D., Robert A. Sweet, M.D., Roger Rosenberg, M.D., Thomas D. Bird, M.D., CC, and Jeremy M. Silverman, Ph.D.

The Alzheimer's Disease Sequencing Project (ADSP) is comprised of two Alzheimer's Disease (AD) genetics consortia and three National Human Genome Research Institute (NHGRI) funded Large Scale Sequencing and Analysis Centers (LSAC). The two AD genetics consortia are the Alzheimer's Disease Genetics Consortium (ADGC) funded by NIA (U01 AG032984), and the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) funded by NIA (R01 AG033193), the National Heart, Lung, and Blood Institute (NHLBI), other National Institute of Health (NIH) institutes and other foreign governmental and non-governmental organizations. The Discovery Phase analysis of sequence data is supported through UF1AG047133 (to Drs. Schellenberg, Farrer, Pericak-Vance, Mayeux, and Haines); U01AG049505 to Dr. Seshadri; U01AG049506 to Dr. Boerwinkle; U01AG049507 to Dr. Wijsman; and U01AG049508 to AG and the Discovery Extension Phase analysis is supported through U01AG052411 to AG, U01AG052410 to Dr. Pericak-Vance and U01 AG052409 to Drs. Seshadri and Fornage. Data generation and harmonization in the Follow-up Phases is supported by U54AG052427 (to Drs. Schellenberg and Wang).

The ADGC cohorts include: Adult Changes in Thought (ACT), the Alzheimer's Disease Centers (ADC), the Chicago Health and Aging Project (CHAP), the Memory and Aging Project (MAP), Mayo Clinic (MAYO), Mayo Parkinson's Disease controls, University of Miami, the Multi-Institutional Research in Alzheimer's Genetic Epidemiology Study (MIRAGE), the National Cell Repository for Alzheimer's Disease (NCRAD), the National Institute on Aging Late Onset Alzheimer's Disease Family Study (NIA-LOAD), the Religious Orders Study (ROS), the Texas Alzheimer's Research and Care Consortium (TARC), Vanderbilt University/Case Western Reserve University (VAN/CWRU), the Washington Heights-Inwood Columbia Aging Project (WHICAP) and the Washington University Sequencing Project (WUSP), the Columbia University Hispanic-Estudio Familiar de Influencia Genetica de Alzheimer (EFIGA), the University of Toronto (UT), and Genetic Differences (GD).

The CHARGE cohorts are supported in part by National Heart, Lung, and Blood Institute (NHLBI) infrastructure grant HL105756 (Psaty), RC2HL102419 (Boerwinkle) and the neurology working group is supported by the National Institute on Aging (NIA) R01 grant AG033193. The CHARGE cohorts participating in the ADSP include the following: Austrian Stroke Prevention Study (ASPS), ASPS-Family study, and the Prospective Dementia Registry-Austria (ASPS/PRODEM-Aus), the Atherosclerosis Risk in Communities (ARIC) Study, the Cardiovascular Health Study (CHS), the Erasmus Rucphen Family Study (ERF), the Framingham Heart Study (FHS), and the Rotterdam Study (RS). ASPS is funded by the Austrian Science Fond (FWF) grant number P20545-P05 and P13180 and the Medical University of Graz. The ASPS-Fam is funded by the Austrian Science Fund (FWF) project I904),the EU Joint Programme-Neurodegenerative Disease Research (JPND) in frame of the BRIDGET project (Austria, Ministry of Science) and the Medical University of Graz and the Steiermärkische Krankenanstalten Gesellschaft. PRODEM-Austria is supported by the Austrian Research Promotion agency (FFG) (Project No. 827462) and by the Austrian National Bank (Anniversary Fund, project 15435. ARIC research is carried out as a collaborative study supported by NHLBI contracts (HHSN268201100005C,HHSN268201100006C,HHSN268201100007C,HHSN268201100008C,HHSN268201100009C,HHSN268201100010C,HHSN268201100011C,andHHSN268201100012C). Neurocognitive data in ARIC is collected by U01 2U01HL096812, 2U01HL096814, 2U01HL096899, 2U01HL096902, 2U01HL096917 from the NIH (NHLBI, NINDS, NIA and NIDCD), and with previous brain MRI examinations funded by R01-HL70825 from the NHLBI. CHS research was supported by contracts

HHSN268201200036C, HHSN268200800007C, N01HC55222, N01HC85079, N01HC85080, N01HC85081, N01HC85082, N01HC85083, N01HC85086, and grants U01HL080295 and U01HL130114 from the NHLBI with additional contribution from the National Institute of Neurological Disorders and Stroke (NINDS). Additional support was provided by R01AG023629, R01AG15928, and R01AG20098 from the NIA. FHS research is supported by NHLBI contracts N01-HC-25195 and HHSN268201500001I. This study was also supported by additional grants from the NIA (R01s AG054076, AG049607 and AG033040 and NINDS (R01 NS017950). The ERF study as a part of EUROSPAN (European Special Populations Research Network) was supported by European Commission FP6 STRP grant number 018947 (LSHG-CT-2006-01947) and also received funding from the European Community's Seventh Framework Programme (FP7/2007-2013)/grant agreement HEALTH-F4-2007-201413 by the European Commission under the programme Quality of Life and Management of the Living Resources of 5th Framework Programme (no. QLG2-CT-2002-01254). High-throughput analysis of the ERF data was supported by a joint grant from the Netherlands Organization for Scientific Research and the Russian Foundation for Basic Research (NWO-RFBR 047.017.043). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, the Netherlands Organization for Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the municipality of Rotterdam. Genetic data sets are also supported by the Netherlands Organization of Scientific Research NWO Investments (175.010.2005.011, 911-03-012), the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), and the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA), project 050-060-810. All studies are grateful to their participants, faculty and staff. The content of these manuscripts is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the U.S. Department of Health and Human Services.

The three LSACs are: the Human Genome Sequencing Center at the Baylor College of Medicine (U54 HG003273), the Broad Institute Genome Center (U54HG003067), and the Washington University Genome Institute (U54HG003079).

Biological samples and associated phenotypic data used in primary data analyses were stored at Study Investigators institutions, and at the National Cell Repository for Alzheimer's Disease (NCRAD, U24AG021886) at Indiana University funded by NIA. Associated Phenotypic Data used in primary and secondary data analyses were provided by Study Investigators, the NIA funded Alzheimer's Disease Centers (ADCs), and the National Alzheimer's Coordinating Center (NACC, U01AG016976) and the National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (NIAGADS, U24AG041689) at the University of Pennsylvania, funded by NIA, and at the Database for Genotypes and Phenotypes (dbGaP) funded by NIH. This research was supported in part by the Intramural Research Program of the National Institutes of health, National Library of Medicine. Contributors to the Genetic Analysis Data included Study Investigators on projects that were individually funded by NIA, and other NIH institutes, and by private U.S. organizations, or foreign governmental or nongovernmental organizations.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2018.00209/full#supplementary-material

Figure S1 | Schematic representation of the bioinformatics pipeline used in this study to filter and combine the information from whole exome sequencing (WES) and whole genome sequencing (WGS).

Table S1 | Structure of the families used in this study with detail of the number of individuals (IDs) sequenced per family, number of cases (CA), number of controls (CO), number of females (Fe) and number of males (Ma).The first 25 families were employed in the simulation analysis.

Table S2 | Design of simulated "GENE-A" across 25 families. Scenarios 5FCx0FNC, 5FCx5FNC, 5FCx10FNC, 5FCx15FNC, 5FCx20FNC.

Table S3 | Comparison of kinship matrices for Fam#1 and Fam#2.

REFERENCES

Aly, T. A., Baschal, E. E., Jahromi, M. M., Fernando, M. S., Babu, S. R., Fingerlin, T. E., et al. (2008). Analysis of single nucleotide polymorphisms identifies major type 1A diabetes locus telomeric of the major histocompatibility complex. *Diabetes* 57, 770–776. doi: 10.2337/db07-0900

Balding, D. J., and Nichols, R. A. (1995). A method for quantifying differentiation between populations at Multi-Allelic Loci and Its implications for investigating identity and paternity. *Genetica* 96, 3–12. doi: 10.1007/BF01441146

Bansal, V., Libiger, O., Torkamani, A., and Schork, N. J. (2010). Statistical analysis strategies for association studies involving rare variants. *Nat. Rev. Genet.* 11, 773–785. doi: 10.1038/nrg2867

Baranzini, S. E., Wang, J., Gibson, R. A., Galwey, N., Naegelin, Y., Barkhof, F., et al. (2009). Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum. Mol. Genet.* 18, 767–778. doi: 10.1093/hmg/ddn388

Beecham, G. W., Hamilton, K., Naj, A. C., Martin, E. R., Huentelman, M., Myers, A. J., et al. (2014). Genome-wide association meta-analysis of neuropathologic features of Alzheimer's disease and related dementias. *PLoS Genet.* 10:e1004606. doi: 10.1371/journal.pgen.1004606

Chen, C., Manichaikul, A., and Rich, S. S. (2009). A generalized family-based association test for dichotomous traits. *Am. J. Hum. Genet.* 85, 364–376. doi: 10.1016/j.ajhg.2009.08.003

Cheong, C., Hentschel, L., Davidson, A. E., Gerrelli, D., Davie, R., Rizzo, R., et al. (2016). Mutations in CPAMD8 cause a unique form of autosomal-recessive anterior segment dysgenesis. *Am. J. Hum. Genet.* 99, 1338–1352. doi: 10.1016/j.ajhg.2016.09.022

Choi, S., Lee, S., Cichon, S., Nöthen, M. M., Lange, C., Park, T., et al. (2014). FARVAT: a family-based rare variant association test. *Bioinformatics* 30, 3197–3205. doi: 10.1093/bioinformatics/btu496

Cingolani, P., Platts, A., Wang le L., Melissa Coon, W., Nguyen, T., Wang, L., et al. (2012). A Program for annotating and predicting the effects of single nucleotide

polymorphisms, SnpEff: SNPs in the genome of drosophila melanogaster strain w1118; Iso-2; Iso-3. *Fly (Austin)* 6, 80–92. doi: 10.4161/fly.19695

Cirulli, E. T., and Goldstein, D. B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nat. Rev. Genet.* 11, 415–425. doi: 10.1038/nrg2779

Cruchaga, C., Celeste,. M. K., Jin, S. C., Benitez, B. A., Cai, Y., Guerreiro, R., et al. (2014). Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease. *Nature* 505, 550–554. doi: 10.1038/nature12825

Cruchaga, C., Del-Aguila, J. L., Saef, B., Black, K., Fernandez, M. V., Budde, J., et al. (2017). Polygenic risk score of sporadic late-onset Alzheimer's disease reveals a shared architecture with the familial and early-onset forms. *Alzheimers Dement.* 14, 205–214. doi: 10.1016/j.jalz.2017.08.013

Cruchaga, C., Haller, G., Chakraverty, S., Mayo, K., Vallania, F. L.M., Mitra, R. D., et al. (2012). Rare variants in APP, PSEN1 and PSEN2 increase risk for AD in Late-onset Alzheimer's disease families. *PLoS ONE* 7:e31039. doi: 10.1371/journal.pone.0031039

Cukier, H. N., Dueker, N. D., Slifer, S. H., Lee, J. M., Whitehead, P. L, Lalanne, E., et al. (2014). Exome sequencing of extended families with autism reveals genes shared across neurodevelopmental and neuropsychiatric disorders. *Mol. Autism* 5:1. doi: 10.1186/2040-2392-5-1

De, G., Yip, W., Ionita-Laza, I., Laird, N., and Amos, C. I. (2013). Rare variant analysis for family-based design. *PLoS ONE* 8:e48495. doi: 10.1371/journal.pone.0048495

Farrer, L. A., Cupples, L. A., Haines, J. L., Hyman, B., Kukull, W. A., Mayeux, R., et. al. (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer disease meta analysis consortium. *JAMA* 278, 1349–1356. doi: 10.1001/jama.1997.03550160069041

Fernández, M. V., Kim, J. H., Budde, J. P., Black, K., Medvedeva, A., Saef, B., et al. (2017). Analysis of neurodegenerative mendelian genes in clinically diagnosed Alzheimer disease. *PLOS Genet.* 13:e1007045. doi: 10.1371/journal.pgen.1007045

Frazer, K. A., Murray, S. S., Schork, N. J., and Topolm, E. J. (2009). Human genetic variation and its contribution to complex traits. *Nat. Rev. Genet.* 10, 241–251. doi: 10.1038/nrg2554

Gaitán-Peñas, H., Apaja, P. M., Arnedo, T., Castellanos, A., Elorza-Vidal, X., Soto, D., et al. (2017). Leukoencephalopathy-Causing CLCN2 mutations are associated with impaired Cl–channel function and trafficking. *J. Physiol.* 595, 6993–7008. doi: 10.1113/JP275087

Guerreiro, R. J., Lohmann, E., Brás, J. M., Gibbs, J. R., Rohrer, J. D., Gurunlian, N., et al. (2013). Using exome sequencing to reveal mutations in TREM2 presenting as a frontotemporal dementia-like syndrome without bone involvement. *JAMA Neurol.* 70, 78–84. doi: 10.1001/jamaneurol.2013.579

He, Z., O'Roak, B. J., Smith, J. D., Wang, G., Hooker, S., Santos-Cortez, R. L. P., et al. (2014). Rare-variant extensions of the transmission disequilibrium test: application to autism exome sequence data. *Am. J. Hum. Genet.* 94, 33–46. doi: 10.1016/j.ajhg.2013.11.021

He, Z., Zhang, D., Renton, A. E., Li, B., Zhao, L., Wang, G. T., et al. (2017). The rare-variant generalized disequilibrium test for association analysis of nuclear and extended pedigrees with application to Alzheimer disease WGS data. *Am. J. Hum. Genet.* 100, 193–204. doi: 10.1016/j.ajhg.2016.12.001

Horvath, S., Xu, X., and Laird, N. M. (2001). The family based association test method: strategies for studying general genotype–phenotype associations. *Eur. J. Hum. Genet.* 9, 301–306. doi: 10.1038/sj.ejhg.5200625

Ionita-Laza, I., Lee, S., Makarov, V., Buxbaum, J. D., and Lin, X. (2013). Family-based association tests for sequence data, and comparisons with population-based association tests. *Eur. J. Hum. Genet.* 21, 1158–1162. doi: 10.1038/ejhg.2012.308

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S., Freimer, N. B., et al. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* 42, 348–354. doi: 10.1038/ng.548

Kazma, R., and Bailey, J. N. (2011). Population-based and family-based designs to analyze rare variants in complex diseases. *Genet. Epidemiol.* 35(Suppl. 1). S41–S47. doi: 10.1002/gepi.20648

Laird, N. M., Horvath, S., and Xu, X. (2000). Implementing a unified approach to family-based tests of association. *Genet. Epidemiol.* 19(Suppl. 1). S36–S42. doi: 10.1002/1098-2272(2000)19:1+ <::AID-GEPI6>3.0.CO;2-M

Laird, N. M., and Lange, C. (2006). Family-based designs in the age of large-scale gene-association studies. *Nat. Rev. Genet.* 7, 385–394. doi: 10.1038/nrg1839

Lambert, J. C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., et al. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* 45, 1452–1458. doi: 10.1038/ng.2802

Lange, C., and Laird, N. M. (2002). On a general class of conditional tests for family-based association studies in genetics: the asymptotic distribution, the conditional power, and optimality considerations. *Genet. Epidemiol.* 23, 165–180. doi: 10.1002/gepi.209

Lee, S., Abecasis, G. R., Boehnke, M., and Lin, X. (2014). Rare-variant association analysis: study designs and statistical tests. *Am. J. Hum. Genet.* 95, 5–23. doi: 10.1016/j.ajhg.2014.06.009

Lee, S., Emond, M. J., Bamshad, M. J., Barnes, K. C., Rieder, M. J., Nickerson, D. A., et al. (2012). Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. *Am. J. Hum. Genet.* 91, 224–237. doi: 10.1016/j.ajhg.2012.06.007

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291. doi: 10.1038/nature19057

Li, B., and Leal, S. M. (2008). Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am. J. Hum. Genet.* 83, 311–321. doi: 10.1016/j.ajhg.2008.06.024

Li, M., Boehnke, M., and Abecasis, G. R. (2006). Efficient study designs for test of genetic association using sibship data and unrelated cases and controls. *Am. J. Hum. Genet.* 78, 778–792. doi: 10.1086/503711

Lin, A., Wang, R. T., Ahn, S., Park, C. C., and Smith, D. J. (2010). A genome-wide map of human genetic interactions inferred from radiation hybrid genotypes. *Genome Res.* 20, 1122–1132. doi: 10.1101/gr.104216.109

Manolio, T. A., Francis, A., Collins, S., Cox, N. J., Goldstein, D. B., and Hindorff, L. A. (2009). Finding the missing heritability of complex diseases. *Nature* 461, 747–753. doi: 10.1038/nature08494

Morgenthaler, S., and Thilly, W. G. (2007). A Strategy to discover genes that carry multi-allelic or mono-allelic risk for common diseases: a Cohort Allelic Sums Test (CAST). *Mutat. Res.* 615, 28–56. doi: 10.1016/j.mrfmmm.2006.09.003

Neale, B. M., Rivas, M. A., Voight, B. F., Altshuler, D., Devlin, B., Orho-Melander, M., et al. (2011). Testing for an unusual distribution of rare variants. *PLoS Genet.* 7:e1001322. doi: 10.1371/journal.pgen.1001322

Neale, B. M., and Sham, P. C. (2004). The future of association studies: gene-based analysis and replication. *Am. J. Hum. Genet.* 75, 353–362. doi: 10.1086/423901

Nyúl-Tóth, Á., Kozma, M., Nagyoszi, P., Nagy, K., Fazakas, C., Haskó, J., et al. (2017). Expression of pattern recognition receptors and activation of the non-canonical inflammasome pathway in brain pericytes. *Brain Behav. Immun.* 64, 220–231. doi: 10.1016/j.bbi.2017.04.010

Ott, J., Kamatani, Y., and Lathrop, M. (2011). Family-based designs for genome-wide Association Studies. *Nat. Rev. Genet.* 12, 465–474. doi: 10.1038/nrg2989

Pontillo, A., Catamo, E., Arosio, B., Mari, D., and Crovella, S. (2012). NALP1/NLRP1 genetic variants are associated with Alzheimer disease. *Alzheimer Dis. Assoc. Disord.* 26, 277–281. doi: 10.1097/WAD.0b013e318231a8ac

Price, A. L., Kryukov, G. V., de Bakker, P. I.W., Purcell, S. M., Staples, J., Wei, LJ., et al. (2010). Pooled association tests for rare variants in exon-resequencing studies. *Am. J. Hum. Genet.* 86, 832–838. doi: 10.1016/j.ajhg.2010.04.005

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: www.R-project.org/

Ridge, P. G., Hoyt, K. B., Boehme, K., Mukherjee, S., Crane, P. K., Haines, J. L., et al. (2016). Assessment of the genetic variance of late-onset Alzheimer's disease. *Neurobiol. Aging* 41, 200.e13–200.e20. doi: 10.1016/j.neurobiolaging.2016.02.024

Rosenthal, S. L, and Kamboh, M. I. (2014). Late-onset Alzheimer's disease genes and the potentially implicated pathways. *Curr. Genet. Med. Rep.* 22, 85–101. doi: 10.1007/s40142-014-0034-x

Saint-Martin, C., Gauvain, G., Teodorescu, G., Gourfinkel-An, I., Fedirko, E., Weber, Y. G., et al. (2009). Two novel CLCN2 mutations accelerating chloride

channel deactivation are associated with idiopathic generalized epilepsy. *Hum. Mutat.* 30, 397–405. doi: 10.1002/humu.20876

Schaid, D. J., McDonnell, S. K., Sinnwell, J. P., and Thibodeau, S. N. (2013). Multiple genetic variant association testing by collapsing and kernel methods with pedigree or population structured data. *Genet. Epidemiol.* 37, 409–418. doi: 10.1002/gepi.21727

Sims, R., van der Lee, S. J., Naj, A. C., Bellenguez, C., Badarinarayan, N., Jakobsdottir, J., 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

Smith, M., Herrell, S., Lusher, M., Lako, L., Simpson, C., and Wiestner, A. (1999). Genomic organisation of the human chordin gene and mutation screening of candidate cornelia de lange syndrome genes. *Hum. Genet.* 105, 104–111. doi: 10.1007/s004399900068

Spielman, R. S., McGinnis, R. E., and Ewens, W. J. (1993). Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). *Am. J. Hum. Genet*. 52, 506–516.

Sul, J. H., Cade, B. E., Cho, M. H., Qiao, D., Silverman, E. K., Redline, S., et al. (2016). Increasing generality and power of rare-variant tests by utilizing extended pedigrees. *Am. J. Hum. Genet.* 99, 846–859. doi: 10.1016/j.ajhg.2016.08.015

Thornton, T., and McPeek, M. S. (2007). Case-control association testing with related individuals: a more powerful quasi-likelihood score test. *Am. J. Hum. Genet.* 81, 321–337. doi: 10.1086/519497

Wang, X., Lee, S., Zhu, X., Redline, S., and Lin, X. (2013). GEE-based SNP set association test for continuous and discrete traits in family-based association studies. *Genet. Epidemiol.* 37, 778–786. doi: 10.1002/gepi.21763

Wang, X., Lopez, O. L., Sweet, R. A., Becker, J. T., Dekosky, S. T., Barmada, M. M., et al. (2015). Genetic determinants of disease progression in Alzheimer's disease. *J. Alzheimers Dis.* 43, 649–655. doi: 10.3233/JAD-140729

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag.

Wijsman, E. M., Pankratz, N. D., Choi, Y., Rothstein, J. H., Faber, K. M., Cheng, R., et al. (2011). Genome-wide association of familial late-onset Alzheimer's disease replicates BIN1 and CLU and nominates CUGBP2 in interaction with APOE. *PLoS Genet.* 7:e1001308. doi: 10.1371/journal.pgen.1001308

Wu, M. C., Lee, S., Cai, T., Li, Y., Boehnke, M., and Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *Am. J. Hum. Genet.* 89, 82–93. doi: 10.1016/j.ajhg.2011.05.029

Yan, Q., Tiwari, H. K., Yi, N., Gao, G., Zhang, K., Lin, W. Y., et al. (2015). A sequence kernel association test for dichotomous traits in family samples under a generalized linear mixed model. *Hum. Hered.* 79, 60–68. doi: 10.1159/000375409

Zhou, X., and Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* 44, 821–824. doi: 10.1038/ng.2310

Zöllner, S., Wen, X., Hanchard, N. A., Herbert, M. A., Ober, C., and Pritchard, J. K. (2004). Evidence for extensive transmission distortion in the human genome. *Am. J. Hum. Genet.* 74, 62–72. doi: 10.1086/381131

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 10 April 2018
doi: 10.3389/fnins.2018.00230

Pleiotropic Effects of Variants in Dementia Genes in Parkinson Disease

Laura Ibanez[1], Umber Dube[1], Albert A. Davis[2], Maria V. Fernandez[1], John Budde[1], Breanna Cooper[1], Monica Diez-Fairen[3,4], Sara Ortega-Cubero[3,5], Pau Pastor[3,4], Joel S. Perlmutter[2,6], Carlos Cruchaga[1†] and Bruno A. Benitez[7†]*

[1] Department of Psychiatry, Washington University, Saint Louis, MO, United States, [2] Department of Neurology, Washington University, Saint Louis, MO, United States, [3] Centro de Investigación Biomédica en Red de Enfermedades Neurodegenerativas, Instituto de Salud Carlos III, Madrid, Spain, [4] Movement Disorders Unit, Department of Neurology, University Hospital Mutua de Terrassa, Fundació per la Recerca Biomèdica i Social Mútua Terrassa, Terrassa, Barcelona, Spain, [5] Department of Neurology and Neurosurgery, Hospital Universitario de Burgos, Burgos, Spain, [6] Departments of Radiology, Neuroscience, Physical Therapy, and Occupational Therapy, Washington University, Saint Louis, MO, United States, [7] Department of Medicine, Washington University, Saint Louis, MO, United States

OPEN ACCESS

Edited by:
Leo P. Sugrue, University of California, San Francisco, United States

Reviewed by:
Maria Shadrina, Institute of Molecular Genetics (RAS), Russia
Jennifer Fifita, Macquarie University, Australia

***Correspondence:**
Bruno A. Benitez
babenitez@wustl.edu

†These authors have contributed equally to this work.

Specialty section:
This article was submitted to Neurodegeneration, a section of the journal Frontiers in Neuroscience

Received: *13 December 2017*
Accepted: *23 March 2018*
Published: *10 April 2018*

Citation:
Ibanez L, Dube U, Davis AA, Fernandez MV, Budde J, Cooper B, Diez-Fairen M, Ortega-Cubero S, Pastor P, Perlmutter JS, Cruchaga C and Benitez BA (2018) Pleiotropic Effects of Variants in Dementia Genes in Parkinson Disease. Front. Neurosci. 12:230. doi: 10.3389/fnins.2018.00230

Background: The prevalence of dementia in Parkinson disease (PD) increases dramatically with advancing age, approaching 80% in patients who survive 20 years with the disease. Increasing evidence suggests clinical, pathological and genetic overlap between Alzheimer disease, dementia with Lewy bodies and frontotemporal dementia with PD. However, the contribution of the dementia-causing genes to PD risk, cognitive impairment and dementia in PD is not fully established.

Objective: To assess the contribution of coding variants in Mendelian dementia-causing genes on the risk of developing PD and the effect on cognitive performance of PD patients.

Methods: We analyzed the coding regions of the amyloid-beta precursor protein (*APP*), Presenilin 1 and 2 (*PSEN1, PSEN2*), and Granulin (*GRN*) genes from 1,374 PD cases and 973 controls using pooled-DNA targeted sequence, human exome-chip and whole-exome sequencing (WES) data by single variant and gene base (SKAT-O and burden tests) analyses. Global cognitive function was assessed using the Mini-Mental State Examination (MMSE) or the Montreal Cognitive Assessment (MoCA). The effect of coding variants in dementia-causing genes on cognitive performance was tested by multiple regression analysis adjusting for gender, disease duration, age at dementia assessment, study site and *APOE* carrier status.

Results: Known AD pathogenic mutations in the *PSEN1* (p.A79V) and *PSEN2* (p.V148I) genes were found in 0.3% of all PD patients. There was a significant burden of rare, likely damaging variants in the *GRN* and *PSEN1* genes in PD patients when compared with frequencies in the European population from the ExAC database. Multiple regression analysis revealed that PD patients carrying rare variants in the *APP*, *PSEN1*, *PSEN2*, and *GRN* genes exhibit lower cognitive tests scores than non-carrier PD patients ($p = 2.0 \times 10^{-4}$), independent of age at PD diagnosis, age at evaluation, *APOE* status or recruitment site.

Conclusions: Pathogenic mutations in the Alzheimer disease-causing genes (*PSEN1* and *PSEN2)* are found in sporadic PD patients. PD patients with cognitive decline carry rare variants in dementia-causing genes. Variants in genes causing Mendelian neurodegenerative diseases exhibit pleiotropic effects.

Keywords: Parkinson disease, dementia, cognitive impairment, rare variants, *APP, PSEN1, PSEN2, GRN*

INTRODUCTION

Recent genome-wide association studies (GWAS) have reported an overlap between Parkinson disease (PD) and the most common forms of dementia including Alzheimer disease (AD), dementia with Lewy bodies (DLB) and Frontotemporal dementia (FTD) (Guerreiro et al., 2015; Ferrari et al., 2017). Multiple variants associated with PD risk have also been identified as risk factors for AD, DLB, or FTD including variants in the following genes: *Triggering receptor expressed on myeloid cells 2 (TREM2), Microtubule-associated protein tau* (*MAPT*), *C9orf72, Glucocerebrosidase* (*GBA)*, and *Apolipoprotein E (APOE)* (Parsian et al., 2002; Harms et al., 2013; Davis et al., 2015; Benitez et al., 2016).

AD is the most common form of dementia and is characterized pathologically by the accumulation of amyloid plaques and neurofibrillary tangles. *Amyloid protein precursor* (*APP*), *presenilin-1* (*PSEN1*), and *presenilin-2* (*PSEN2*) mutations cause autosomal dominant forms of early-onset AD (Cruts et al., 2012). Approximately 220 pathogenic mutations in the *PSEN1* gene have been reported in AD patients worldwide, whereas 27 and 16 pathogenic mutations have been described in the *APP* and *PSEN2* genes, respectively (Cruts et al., 2012). Rare functional variants in *APP* (Schulte et al., 2015) and pathogenic mutations in *PSEN1*, and *PSEN2* have been also reported in PD patients (Takao et al., 2002; Jimenez-Escrig et al., 2004; Puschmann et al., 2009; Niwa et al., 2013). We recently reported the presence of *leucine-rich repeat kinase 2* (*LRRK2),* p.G2019S mutation in members of two multigenerational families with AD and a suggestive association of variants in the *PTEN-induced putative kinase 1* (*PINK1)* gene with AD (Fernández et al., 2017). These results suggest a genetic overlap between familial AD and PD.

There is also an overlap of neuropathology in the brains of AD and PD patients. Approximately 50–60% of the sporadic and familial AD patients with pathogenic mutations in *APP*, *PSEN1*, or *PSEN2* genes exhibit widespread α-synuclein pathology (Meeus et al., 2012a). Abnormal cortical amyloid-beta (Aβ) deposition is present in 60% of PD patients with dementia (Kotzbauer et al., 2012), the burden of Aβ plaques inversely correlates with cognitive status in PD cases with dementia (Irwin et al., 2013) and the progression of the dementia in PD correlates with Lewy body (LB) and cortical AD-type pathology (Compta et al., 2011). In addition, AD-like changes in cerebrospinal fluid (CSF) biomarkers (Aβ levels) have been reported in PD patients (Terrelonge et al., 2015). Taken together, these results suggest that abnormal APP processing and Aβ accumulation occurs in PD. Here, we evaluate the genetic variation of *APP*, *PSEN1* and *PSEN2* genes in PD patients.

FTD may cause up to 10% of all cases of dementia and is the second most common cause of early-onset dementia (<65 years of age) (Ratnavalli et al., 2002). The three most common genetic causes of FTD are mutations in the genes *MAPT* and *granulin* (*GRN*), and expansions of a hexanucleotide repeat (GGGGCC) in the *C9orf72* gene. *C9orf72* expansions cause 5–12% of all FTD and 10–35% of familial FTD. *MAPT* or *GRN* mutations are found in 1–10% of all FTD and 5–25% of familial FTD. *MAPT* variants associated with FTD also increase the risk of developing PD (Pastor et al., 2000; Benitez et al., 2016). *C9orf72* expansions have been found in PD patients in some (Baizabal-Carvallo and Jankovic, 2016) but not all studies (Harms et al., 2013). Parkinsonism precedes the cognitive and behavioral symptoms of FTD by several years in patients with mutations in *GRN* (Baizabal-Carvallo and Jankovic, 2016). Up to 41% of FTD patients with *GRN* mutations exhibit parkinsonism (Josephs et al., 2007). In addition, *GRN* pathogenic mutations have been reported in PD patients (Brouwers et al., 2007; Rovelet-Lecrux et al., 2008). Here, we explore the role of variants in the *GRN* gene in PD patients.

PD is the most common neurodegenerative movement disorder, affecting ∼1–2% of people over 60 years of age (Wright Willis et al., 2010). Clinically, PD patients exhibit bradykinesia, rest tremor, rigidity, and disturbances in balance. PD is characterized neuropathologically by the presence of α-synuclein-positive neuronal inclusions, commonly referred to as LBs and Lewy neurites, as well as neuronal loss in the substantia nigra. Genetically, familial autosomal dominant PD is caused by highly penetrant mutations in the *alpha-synuclein (SNCA)* and *LRRK2* genes, whereas autosomal recessive PD is caused by mutations in the *PARK2/PARKIN*, *PINK-1*, and *PARK7/DJ-1* genes (Petrucci et al., 2014). Recently, a GWAS identified at least 41 loci associated at genome-wide significant level with disease risk in individuals of European ancestry (Chang et al., 2017). The most statistically significant signals associated with PD are common variants located close to *SNCA*, *MAPT*, and *GBA* genes (Chang et al., 2017).

Dementia is one of the most common non-motor symptoms in PD (Foltynie et al., 2004; Emre et al., 2007). The prevalence of dementia in PD patients at any stage of disease ranges from 22 to 48% (Foltynie et al., 2004; Emre et al., 2007). Several risk factors for dementia in PD include severity of parkinsonism, the presence of non-motor symptoms, older age, male sex and presence of cognitive symptoms at PD diagnosis (Aarsland and Kurz, 2010). Relatives of PD patients exhibit a higher risk of dementia than relatives of control subjects,

(Rocca et al., 2007) supporting a role of genetic factors in the development of dementia in PD. However, the genetic contributors to dementia in PD have yet to be fully established. Longitudinal studies have shown that the *APOE* ε4 allele, the *H1* haplotype in *MAPT* and mutations in the *GBA* gene are associated with a more rapid cognitive decline in PD (Morley et al., 2012; Collins and Williams-Gray, 2016; Liu et al., 2016). In contrast, a cross-sectional study showed that PD carriers of the *LRRK2,*p.G2019S mutation exhibit lower rates of dementia (Srivatsal et al., 2015). However, it is not clear how frequently mutations in dementia-causing genes occur in PD patients or whether variants in these genes relate to dementia in PD patients.

Here, we performed a systematic screening of known dementia-causing genes (*APP*, *PSEN1, PSEN2*, or *GRN* genes) in 821 PD cases and 423 controls from North America in addition to 553 PD patients and 550 healthy controls from Spain.

MATERIALS AND METHODS

Subjects

Three cohorts were included in this study: the Washington University in Saint Louis (WUSTL) Movement Disorder Center (MO, USA), the Parkinson's Progression Markers Initiative (PPMI) consortium (www.ppmi-info.org) and the Movement Disorders Unit at the University of Navarra (UN), School of Medicine (Navarra, Spain). PD clinical diagnoses were based on UK Brain Bank criteria (Hughes et al., 1992). Written informed consent was obtained from all participants prior to their enrollment. The Washington University in Saint Louis Human Research Protection Office (approval number: 201107095) approved the study. Demographic characteristics of these three populations have been published previously (Weintraub et al., 2013; Davis et al., 2015; Benitez et al., 2016). *The WUSTL cohort* included 490 non-hispanic white (NHW) PD cases [64% males, mean $\pm$ SD age at onset (AAO) 60 $\pm$ 11 years and 10% had family history of PD] and age- and population-matched 289 controls (64.8 $\pm$ 10.2 years; mean $\pm$ SD). *The PPMI cohort* was composed of 331 NHW PD cases (50% males, AAO 61 $\pm$ 11 years, 9% had family history of PD) and 134 age- and population-matched controls (60.9 $\pm$ 11.4 years). Finally, *the UN cohort* was composed of 553 Spanish PD cases (59% males, AAO 60 $\pm$ 9 years, 20% had family history of PD) and 550 healthy Spanish age-matched controls (62 $\pm$ 7 years). Only one member of each family with PD was included in the analyses. All individuals carrying pathogenic mutations in *LRRK2, p.G2019S (8 PD cases WUSLT)*, *PARK2, p.D53X (1 PD case WUSTL), or PINK, R492X1(1 PD case WUSTL)* genes, duplications in the *SNCA* gene *(1 PD case WUSTL)* or risk-associated variants in the *TREM2, p.R47H (4 WUSTL)*, *GBA, p.N370S (7 WUSTL),* or *MAPT, p.A152T (4 WUSTL)* genes (Benitez and Cruchaga, 2013; Benitez et al., 2016) were excluded from this study. Principal component analyses (PCA) was conducted to infer the genetic structure of individuals using the EIGENSTRAT software (Price et al., 2006). Only subjects that clustered with PCs of NHW origin in the WUSTL or PPMI cohorts were included in this study.

Sequencing Methods

WUSTL and UN samples: To screen for novel variants, pooled-DNA sequencing was performed, as described previously (Jin et al., 2012; Benitez et al., 2016). Briefly, 62 PCR reactions that covered 46,319 bases of the four selected dementia genes were performed in two equimolar pools of 114 and 98 samples. After ligation, concatenated PCR products were randomly sheared by sonication and prepared for sequencing on an Illumina Genome Analyzer IIx (GAIIx) according to the manufacturer's specifications. The resulting reads were re-aligned to the human genome reference assembly build 36.1 (hg18) using SPLINTER. pCMV6-XL5 amplicon (1908 base pairs) was included in the reaction as a negative control. As positive controls, 10 different constructs (p53 gene) with synthetically engineered mutations at a relative frequency of one mutated copy per 200 normal copies was amplified and pooled with the PCR products. Six human DNA samples heterozygous for previously known mutations in *GRN* and *PSEN1*genes were also included (Jin et al., 2012; Benitez et al., 2016). SPLINTER uses the positive control to estimate sensitivity and specificity for variant calling. The wild-type: mutant ratio in the positive control is similar to the relative frequency expected for a single mutation in one pool (1 chromosome mutated in 100 samples = 1/200). SPLINTER uses the negative control (first 900 bp) to model the errors across the 36-bp Illumina reads and to create an error model from each sequencing run of the machine. Based on the error model SPLINTER calculates a *p*-value for the probability that a predicted variant is a true positive. A *p*-value at which all mutants in the positive controls were identified was defined as the cut-off value for the best sensitivity and specificity. All mutants included as part of the amplified positive control vector were found upon achieving ~30-fold coverage at mutated sites (sensitivity= 100%) and ~80 sites in the 1908 bp negative control vector were predicted to be polymorphic (specificity = ~95%). The variants with a *p*-value below this cut-off value were considered for follow-up confirmation. An average coverage of 30X-fold per haploid genome per pool is the minimum coverage necessary to get an optimal positive predictive value for the SNP-calling algorithm (Vallania et al., 2010). Supplementary Table 1 contains a summary of exon coverage per gene. On average, the coverage was 59.5x per allele per individual, which translate to a total coverage of ~13,566x or ~11,662x depending on the pool size. All evaluated variants were validated by genotyping with Sequenom iPLEX or KASPar techniques in all samples. Common variants (>5%) or synonymous variants were not followed up. *APOE* genotype was obtained by direct genotyping of rs7412 and rs429358 using Taqman technology (Cruchaga et al., 2012b). Most of rare variants were also validated by genome-wide data generated with the NeuroX custom chip in the WUSTL samples. NeuroX chip includes both the standard Illumina exome content ($\approx$ 240,000 variants) and over 24,000 variants associated with neurologic diseases (Nalls et al., 2015). *PPMI Sample:* The VCF files containing the whole exome sequence for all PPMI samples were downloaded from the PPMI website (www.ppmi-info.org). The regions of interest for dementia-related genes were extracted from those files for further analysis. *APOE* genotype (rs7412 and rs429358) was obtained from the VCF files.

Bioinformatics and Statistical Analyses

Annotation and Variant Selection

Variants were annotated using SeattleSeq Annotation (Ng et al., 2009), the Exome Variant Server (EVS), (http://evs.gs.washington.edu/EVS/ - Release ESP6500SI-V2), the ExAC data (release 0.3.1) (Lek et al., 2016) and the Ensembl Genome browser (Aken et al., 2016). Predictions of variant pathogenicity were based on the Genomic Evolutionary Rate Profiling (GERP) (Davydov et al., 2010) and the Combined Annotation Dependent Depletion (CADD) algorithm (v1.3, http://cadd.gs.washington.edu) (Kircher et al., 2014). All putative damaging variants (scores GERP $\geq$ 2.95 and CADD $\geq$ 12.37) (Amendola et al., 2015) were further cross-referenced with The AD&FTD mutation database (Cruts et al., 2012) and ClinVar (Landrum et al., 2016) to identify those previously established with pathogenicity for AD or FTD. Only variants annotated or predicted to be coding variants [missense, nonsense (stop/start gain/loss) and frameshift variants] were included in the analyses. Synonymous variants were not included in our analyses. This analysis was applied independently to all three cohorts.

Single Variant Analysis

Association analysis between PD cases and in-house controls was performed using logistic regression, assuming an additive model for allelic effects. The analysis was adjusted for age, gender, and PCs (using the first two principal components) to correct for potential population stratification using Plink1.9 (Purcell et al., 2007; Chang et al., 2015). The minor allele frequency (MAF) of each variant in PD cases was also compared with the MAF described in the Non-Finnish European subgroup (NFE) from ExAC (Lek et al., 2016). Only coding regions with high-quality (PASS filter) variants reported in ExAC were included in our analyses.

Gene Based Analysis

The gene-based association in PD cases and in-house controls was performed using SKAT-O, which utilizes the R package SKAT (Wu et al., 2011). In addition, the burden of rare protein-altering variants in PD cases from the WUSTL and PPMI cohorts were compared with the burden observed in NFE samples from ExAc by collapsing the counts of all missense, nonsense and frameshift variants in each gene with a MAF < 0.01 and then, Fisher's exact test with Yates correction was applied. ExAC cannot be regarded as a pure control dataset. However, several studies have used this resource as a proxy for reassessing the effect of rare variants in Mendelian genes in the general population in different complex diseases (Roberts et al., 2015; Walsh et al., 2017).

Cognitive Impairment Assessment and Analysis

Cognitive impairment was assessed by the Folstein Mini-Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA) in the WUSTL and PPMI datasets respectively. For the combined analyses, z scores were derived by converting the mean raw MoCA and MMSE scores and standard deviation (SD) to the standard normal distribution with mean 0 and SD 1. Multiple regression analyses were performed in the residuals adjusting by age at diagnosis, age at dementia assessment, gender, study site and *APOE* carrier status as covariates (PROC GLM, SAS). All validated variants were included in the model independently of their clinical interpretation. Global cognitive impairment was defined by an MMSE $\leq$ 25 according to the recommendation the International Parkinson and Movement Disorder Society (MDS) Task Force (Dubois et al., 2007). The recommended MoCA cut-off for PD of $<$26 was applied (Dalrymple-Alford et al., 2010). Patients with a diagnosis of dementia [PPMI ($N = 5$) and WUSTL ($N = 9$)] were excluded from the analyses to avoid disproportionate leverage on the statistical models (Thaler et al., 2012). Nominally significant p-value threshold was set at 0.05. Multiple-test correction cutoff for the single-variant analysis using Bonferroni correction for 4 tests is 1.3×10^{-2}. All statistical tests are two-sided unless otherwise stated, and were performed using Statistical Analysis System (SAS Institute Inc) or GraphPad Prism 5.0.

APOE Analysis

The effect of *APOE* allele and genotype in PD risk was tested comparing the frequency in PD cases and controls. Comparisons were made using the X^2 method. Multiple regression adjusting by age at diagnosis, age at dementia assessment and gender was performed to evaluate the effect of *APOE* ϵ4 allele on MoCA or MMSE scores (PROC GLM, SAS).

RESULTS

Single Variant Analyses

WUSTL Cohort

Twenty-one coding variants were found among the screened genes (Supplementary Table 2). 14.3% (3) variants were novel, 4.8% (1) were reported as known pathogenic variants, 19% (4) were classified as pathogenic nature unclear and 33% (7) were reported as non-pathogenic. The three novel variants were *PSEN1* (p.P303L), *PSEN2* (p.C358R), and *GRN* (p.A29V). Both the *PSEN1* and *PSEN2* variants were found in late-onset PD individuals whereas the *GRN* variant was found in one control (Supplementary Table 3A). The *PSEN1* p.A79V mutation (a known AD pathogenic mutation) was found in three PD cases: An early-onset PD case (44 years old at onset) and two late-onset PD cases (75 and 64 years old at onset, respectively). None of the three carriers reported PD or AD family history. Neurological evaluation of these three patients at an age of 46, 86, and 82 years revealed no evidence of dementia after 2, 11, and 18 years of disease onset, respectively. The carrier of the "probable pathogenic" PSEN2, p.S130L variant exhibited an AAO of 45 years and rapid progression of cognitive impairment (Supplementary Table 3B).

PPMI Cohort

Eighteen coding variants were found in 5% of the cases and 8.9% of the controls (Supplementary Table 4). 22% (4) of the variants were novel, 28% (5) were classified as pathogenic nature unclear and 28% (5) were reported as non-pathogenic; 22% (4) were reported in public databases but with an unknown clinical significance. The four variants of unknown significance included *APP* (p.R499C), *APP* (p.R397T), *APP* (p.Q138R),

and *GRN* (p.C260R). Both *APP* p.R499C and p.R397T were found in late-onset PD (LOPD; AAO >50 years) patients whereas *APP* (p.Q138R) and *GRN* (p.C260R) were found in controls (Supplementary Table 5). Five variants in the *PSEN2* gene were considered of pathogenic nature unclear for AD: p.R62H, p.R71W, p.S130L, p.K161R, and p.V393M. Two LOPD patients carried *PSEN2* p.S130L variant and were cognitively normal at the last clinical assessment (Supplementary Table 5). Interestingly, the *PSEN2* p.R71W variant was present in five controls and two PD patients achieving nominal significance ($p = 0.01$; $OR = 0.16$; 95% $CI = 0.02$–0.98).

UN Cohort

Ten coding variants were found in 4.9% of cases from the Spanish PD cohort. No novel variants were found (Supplementary Table 6). Nine variants were present in the AD/FTD mutation database. 50% (5) of them were considered non-pathogenic, 30% (3) of them were of pathogenic nature unclear, and 10% (1) of them (*PSEN2* p.V148I) was considered to be pathogenic (Cruts et al., 2012). The *PSEN2* p.V148I carrier is an early onset PD case (25 years at onset) with a tremor-dominant parkinsonism and positive family history of PD but dementia-free at last assessment after 12 years of PD. Seven (1.3%) carriers of the *PSEN2* p.S130L variant were found in Spanish PD patients. Three had positive PD family history, three had EOPD, three had psychiatric co-morbidities and two with dementia PD (Supplementary Table 7).

Gene-Based Analysis

The gene-based analysis is a powerful tool to uncover genetic association. In previous studies, we used gene-based analysis of the *GBA* gene in PD and *TREM2* in AD to identify additional variants associated with risk (Jin et al., 2015; Benitez et al., 2016). SKAT-O analysis revealed that none of the dementia-related genes achieved statistical significance in the WUSTL series *[APP* ($p = 0.89$), *GRN* ($p = 0.63$), *PSEN1* ($p = 0.13$), and *PSEN2* ($p = 0.5$)] or PPMI [*APP* ($p = 0.15$), *GRN* ($p = 0.6$), *PSEN1* ($p = 0.9$) and *PSEN2* ($p = 0.09$)] compared with in-house controls (Supplementary Tables 2, 4). However, joint burden analysis revealed a significant enrichment of rare variants in the *GRN* (6.6×10^{-03}; $OR = 1.9$; 95% $CI = 1.2$–3.0) and *PSEN1* ($p = 9.2 \times 10^{-41}$; $OR = 54.2$; 95% $CI = 18.8$–156.1) genes in PD cases compared with the ExAC NFE cohort (**Table 1**). The association of *PSEN1* and PD was maintained ($p = 4.9 \times 10^{-66}$) after excluding the *PSEN1* p.A79V mutation from the analysis.

APOE Association With Status and Cognitive Test Performance

APOE has previously been associated with cognitive impairment in PD patients (Parsian et al., 2002; Morley et al., 2012; Tsuang et al., 2013) but APOE effect on PD risk is still controversial (Federoff et al., 2012). Here, no association was found between different *APOE* alleles (ε4 or ε2) and PD case-control status (Supplementary Tables 8A,B). The *APOE* ε4 allele was not associated with lower MoCA scores among PD patients in the PPMI cohort ($p = 0.56$). However, consistent with a previous report, the presence of the *APOE* ε4 allele was associated with a lower MMSE score ($p = 9.0 \times 10^{-3}$) in WUSTL PD patients (Parsian et al., 2002).

TABLE 1 | Enrichment of rare variants in the *PSEN1* and *GRN* genes in PD patients.

Gene	cMAF* PD Cases	cMAF ExAC$^{\varphi}$	*p*-value$^{\#}$	OR (95% CI)$^{\#}$
APP	0.0004	0.0004	ns§	-
GRN	0.0010	0.0005	6.6×10^{-03}	1.9 (1.2–3.0)
PSEN1	0.0010	0.00002	9.1×10^{-41}	54.2 (18.8–156.1)
PSEN2	0.0010	0.0010	ns	-

**cMAF= cumulative minor allele frequency of all non-synonymous variants in each gene.*
$^{\varphi}$*Non-Finnish European ExAC individuals.*
$^{\#}$*WUSTL and PPMI PD cases vs ExAC non-Finish Europeans controls.*
§*Not statistically significant.*

Effect of Variants in Dementia-Causing Genes on Cognitive Tests Performance

Several rare variants in the *GBA* gene have been associated with lower MMSE scores in PD patients (Liu et al., 2016). We hypothesized that variants in the dementia-causing genes would affect performance on cognitive tests. Interestingly, 4.9% of all PD patients in the PPMI cohort carrying rare variants in dementia-causing genes exhibit significantly ($p = 3.0 \times 10^{-2}$) lower MoCA scores than non-carriers PD patients (**Figure 1A**). Similarly, PD patients carrying rare variants in dementia-causing genes (7.9% of all patients) exhibit significantly ($p = 2.0 \times 10^{-3}$) lower MMSE scores than non-carriers in the WUSTL cohort (**Figure 1B**). Cognitive impairment was assessed with a different cognitive test in each cohort included in this study (MMSE for WUSTL and MoCA for PPMI). Therefore, we perform a combined analysis with age at PD diagnosis, age at which the dementia test was performed, *APOE* status, and cohort included as covariates. PD patients carrying rare variants in dementia-causing genes exhibited lower scores on cognitive tests than non-carrier PD patients ($p = 2.0 \times 10^{-4}$), independent of age at PD diagnosis, age at evaluation, *APOE* status or cohort (**Table 2**).

DISCUSSION

The genetic architecture of dementia in PD has not yet been fully established. Genetic variants in *APOE* and *MAPT* have time-dependent effects on cognition, which vary with disease stage: *MAPT* appears to have its greatest impact on cognitive decline in early PD, whereas *APOE* may have a more pronounced effect late in the course of the disease (Collins and Williams-Gray, 2016). A recent GWAS on an extensive neuropsychological battery in PD patients failed to replicate prior associations with *APOE*, *MAPT*, *catechol-O-methyltransferase (COMT)*, or *SNCA* (Mata et al., 2017). In cross-sectional studies, the prevalence of dementia in *GBA*-PD cases is about 50%, compared to 24–31% in idiopathic PD cases (Setó-Salvia et al., 2012). Longitudinal studies have confirmed a faster progression to dementia in PD cases carrying *GBA* mutations compared to idiopathic PD in a population-representative cohort followed for ∼10 years from diagnosis (Winder-Rhodes et al., 2013). However, *GBA* variants

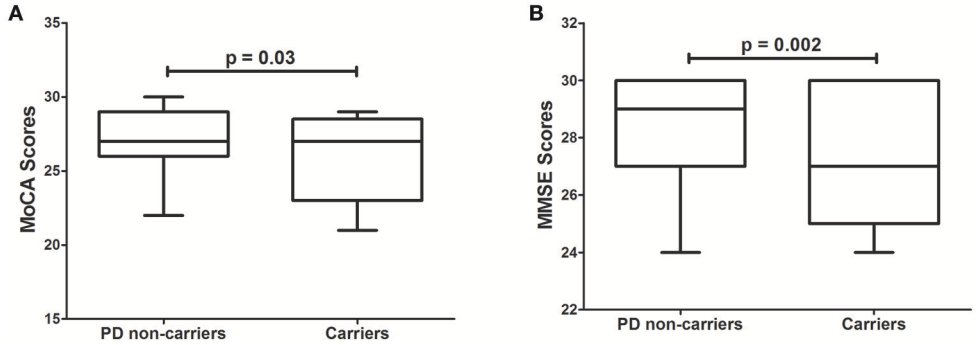

FIGURE 1 | Effect of variants in *APP*, *GRN*, *PSEN1*, and *PSEN2* genes on cognitive test scores in PD patients. **(A),** Tukey's boxplot of MoCA scores in PD non-carriers vs. PD carriers of variants in the dementia-causing genes in the PPMI cohort. **(B),** Tukey's boxplot of the MMSE scores in non-carriers and carriers of variants in the dementia-causing genes in the PD cases of the WUSTL cohort. Student's *t*-test was used and *p*-values reported are based on two-tailed comparisons.

TABLE 2 | Effect of rare variants in dementia-causing genes on cognitive test scores in PD patients.

Outcome	Independent variable	Estimate	Standard error	*t*-value	*p*-value
Cognitive test score$^{\varphi}$	Age at PD Dx§	0.006	0.009	0.64	0.5225
	Age at assessment	−0.015	0.009	−1.64	0.1018
	APOE*	−0.039	0.041	−0.96	0.3385
	Carrier status$^{\#}$	−0.684	0.180	−3.80	0.0002
	Recruitment sites$^{\varphi}$	0.223	0.142	1.58	0.1153

§*Dx = diagnostic.*
$^{\varphi}$*Combined analysis MoCA and MMSE.*
**APOE genotype 22 = 0, 23 = 1, 33 = 2, 24 = 3, 34 = 4, and 44 = 5.*
$^{\#}$*0 = non carrier, 1 = carrier.*
$^{\varphi}$*0 = PPMI, 1 = WUSTL.*

explained only up to 1.4% of PD patients with cognitive decline (Liu et al., 2016). Our results show that 6.7% of PD patients carry rare coding variants in dementia-causing genes and exhibit lower scores on cognitive tests in two independent cohorts compared to non-carriers. The risk of developing dementia varies according to the duration of PD and age at onset (Rocca et al., 2007). However, the results of a covariate-adjusted model confirm that the association between rare variants in dementia-causing genes and lower scores in cognitive tests appears to be independent of age, cohort, disease duration, or *APOE* status.

The *APOE* ε4 allele has been associated with a higher prevalence of dementia in PD (Morley et al., 2012; Tsuang et al., 2013). The small number of PD cases with dementia in each study, the significant heterogeneity of odds ratios between studies, and evidence of publication bias limits the confidence of the *APOE* and dementia in PD association (Huang et al., 2006). In this study, the *APOE* ε4 allele was not associated with risk of developing PD in none of the cohorts (Supplementary Tables 8A,B). However, in the WUSTL cohort, *APOE* ε4 carriers exhibited lower scores on cognitive tests, but the PPMI cohort did not replicate these findings. Therefore, further studies are needed to clarify the role of *APOE* in PD cognitive impairment (Parsian et al., 2002).

The wide variation reported in the prevalence of cognitive impairment (CI) in PD across studies may be due to the cognitive tests employed (Goldman and Litvan, 2011; Burdick et al., 2014). The MMSE and MoCA are the most commonly used tests to assess CI in PD. However, controversy remains regarding the sensitivity of MMSE in assessing cognition in PD patients (Burdick et al., 2014). Even though the MoCA is more sensitive for detecting cognitive changes in PD patients (Hoops et al., 2009), the MMSE was the cognitive test that showed the effects of *GBA* and *LRRK2* mutations on CI in PD patients (Srivatsal et al., 2015; Liu et al., 2016) and is the cognitive test recommended by the International Parkinson and Movement Disorder Society Task Force (Dubois et al., 2007). Nevertheless, here we show that rare variants in dementia-related genes affect both MMSE and MoCA scores in two heterogeneous PD populations ($p = 2.0 \times 10^{-4}$; **Table 2**).

In addition, 0.3% of all PD patients screened in this study carry known AD pathogenic variants in *PSEN1* and *PSEN2* genes (Supplementary Tables 2, 6). A frequency of 0.3% is comparable to the percentage of PD patients carrying the most common pathogenic mutation (*LRRK2* p.G2019S) known for sporadic PD (0.4–1%) (Healy et al., 2008; Correia Guedes et al., 2010). None of the PD patients carrying pathogenic mutations in *PSEN1* and *PSEN2* exhibited signs of dementia at their last clinical assessment. Unfortunately, CSF biomarkers or Pittsburgh compound B (PiB) imaging were not available from these carriers to help determine their preclinical status. The *PSEN1* p.A79V mutation is associated with a broad range of AAO in AD and DLB patients (Cruchaga et al., 2012a; Meeus et al., 2012b). *PSEN1* p.A79V is found in non-demented carriers at an AAO as late as 78 years (Kauwe et al., 2007) and, in some multigenerational AD families, it does not segregate perfectly with disease status (Cruchaga et al., 2012a). However, the *PSEN1* p.A79V mutation increases the Aβ42 level and Aβ42/Aβ40 ratio *in vitro* (Kumar-Singh et al., 2006; Kauwe et al., 2007) but unlike other *PSEN1*

mutations that cause massive Aβ42 deposition, the average Aβ42 brain deposition in sporadic AD patients exceeded those observed in the brains of p.A79V carriers (Kumar-Singh et al., 2006). In addition, non-demented p.A79V carriers also exhibit changes in CSF Aβ42 levels and the Aβ42/40 ratio with no evidence of β-amyloid deposition using PiB imaging (Kauwe et al., 2007). Thus, *PSEN1* p.A79V may be acting through a mechanism other than increasing Aβ deposition. Additional mutations in the *PSEN1* gene including p.G217D, p.V272A, and p.L420R have been associated with dementia and parkinsonism (Takao et al., 2002; Jimenez-Escrig et al., 2004; Puschmann et al., 2009; Niwa et al., 2013). Recently, it was reported that a *de novo PSEN1* mutation is responsible for an early-onset parkinsonism with cognitive impairment (Carecchio et al., 2017). In addition to the clinical overlap between dementia and parkinsonism in carriers of *PSEN1* mutations, there is extensive Lewy body pathology in early-onset AD carriers of the *PSEN1* p.S170F (Snider et al., 2005) and *PSEN1* p.A431V mutations (Qiao et al., 2017) suggesting an interaction between *PSEN1* dysfunction and α-synuclein aggregation.

We also found the *PSEN2* p.V148I mutation in a Spanish PD patient but not in 550 age-matched Spanish controls. The *PSEN2* p.V148I mutation was originally reported in a Spanish patient with late-onset AD (AAO = 71 years) (Lao et al., 1998) but, its pathogenicity has been questioned based on the absence of effect on either Aβ42 levels, Aβ40 levels, or the Aβ42/40 ratio *in vitro* (Walker et al., 2005). However, some variants that show no effect on Aβ42 levels or the Aβ42/Aβ40 ratio (Walker et al., 2005) affect calcium signaling in cultured skin fibroblasts from mutation carriers (Li et al., 2006). Increasing evidence suggests a role of the *PSEN2* p.S130L variant in PD and dementia. Here, we found a total of nine PD patients and one control carrying the "probably pathogenic" *PSEN2* p.S130L variant. The *PSEN2* p.S130L variant has been reported in patients with late-onset AD and mild bradykinesia (Tomaino et al., 2007; Lohmann et al., 2012). Two siblings of a *PSEN2* p.S130L carrier AD patient were diagnosed with PD (Tomaino et al., 2007). Recently, *PSEN2* p.S130L was reported in an individual with idiopathic PD with dementia (AAO = 73 years) (Schulte et al., 2015). Another mutation in *PSEN2* (p.V191E) also was found in one late-onset PD patient (AAO = 75 years) with cognitive decline (Meeus et al., 2012b). Interestingly, a Swedish PD family who carry a *de novo* α-synuclein p.A53T mutation also carried the *PSEN2* p.R163H variant (Puschmann et al., 2009). Carriers of both mutations develop early-onset dementia (Puschmann et al., 2009). All these findings support our data that pathogenic variants in the presenilin genes are present in a small proportion of sporadic PD patients and contribute to α-synuclein aggregation.

We found an enrichment of rare variants in the *GRN* gene in PD patients compared to the general population (ExAC NFE). The frequencies of the p.R433W and p.R478H variants were higher in the WUSTL cohort than in the NFE ExAC cohort. Interestingly, the variant p.R433W was reported in neuropathologically confirmed LBD cases (Meeus et al., 2012b). Mutations in the *GRN* gene occur in LBD patients (Meeus et al., 2012b). A heterozygous deletion removing exons 1 to 11 of the *GRN* gene was reported in an 83 year-old PD patient (Rovelet-Lecrux et al., 2008) and, the IVS0 + 5G>C mutation was reported in a 56 year-old PD patient (Brouwers et al., 2007). In addition, parkinsonism occurs in some FTD patients and is more common in those patients (up to 41%) with *GRN* haploinsufficiency (Josephs et al., 2007). Interestingly, *GRN* overexpression in the substantia nigra protected nigrostriatal neurons in a mouse model of PD (Van Kampen et al., 2014). *GRN* seems to play an important role in multiple neurodegenerative diseases including PD, likely due to its function as a neurotrophic factor and its recently uncovered lysosome function (Tanaka et al., 2013).

Low frequency and rare mutations in the *GBA* gene, which encodes the lysosomal enzyme β-glucocerebrosidase-1, consistently relate to CI in PD and to Lewy body dementia (Nalls et al., 2013). α-synuclein is mainly degraded by lysosomes (Cuervo et al., 2004) and lysosomal dysfunction may contribute to *de novo* aggregation of α-synuclein and impaired autophagic degradation of cytosolic aggregates (Bourdenx et al., 2014). LBs and Lewy neurites may seed around impaired lysosomes and grow in size by continuous deposition of lysosomal-derived un-degraded material as the disease progresses (Dehay et al., 2013). Thus, considering that presenilin and granulin are lysosomal proteins (Sannerud et al., 2016; Kao et al., 2017), it is logical to suggest that variants in the presenilin or granulin genes may exacerbate the cognitive impairments in PD by affecting lysosomal function and facilitating cell-to-cell transfer of proteopathic seeds in the progression of synucleinopathies.

The relatively small size of this study limits the statistical power, which could be the reason we failed to find significant associations between the PD cases and the in-house controls. However, the inclusion of the large NFE ExAC cohort with similar genetic background minimized this limitation. The data used in this study were not generated using a single sequencing method. However, none of the approaches used were expected to have 100% sensitivity for variant detection. Although these technical limitations could have marginal effects on estimates of rare variant frequency and odds ratio values, we do not expect them to alter the key conclusions of this study. Further studies are needed to confirm the role of variants in dementia genes in the cognitive impairment found in PD patients.

CONCLUSION

Our study shows that rare variants in several dementia-related genes are enriched in PD patients compared with normal controls. The PD patients with these variants exhibited lower cognitive performance than PD patients without these variants. Moreover, known dementia-causing mutations are found in PD patients.

AUTHOR CONTRIBUTIONS

BB and CC conceived and designed the study. JP, AD, UD, SO-C, and PP acquired and analyzed the clinical data. LI, BB, MF, JB, and MD-F acquired the genetic data. LI, BB, and CC performed the statistical analysis and interpreted the genetic data. LI and

BB wrote the draft of the manuscript and JP, AD, PP, and CC provided critical comments on the draft of the manuscript. All authors read and approved the final version of the manuscript.

FUNDING

This work was supported by grants from NINDS/NIA (NS075321, NS41509, NS058714, and R01-AG035083); the Barnes-Jewish Hospital Foundation (BJHF); the American Parkinson Disease Association (APDA) Advanced Research Center for Parkinson Disease at Washington University in St. Louis; the Greater St. Louis Chapter of the APDA; the Barnes-Jewish Hospital Foundation (Elliot Stein Family Fund and Parkinson Disease Research Fund), The Michael J. Fox Foundation for Parkinson's Research, Alzheimer's Association and Weston Brain Institute (BAND-14-338165). This research was conducted while CC was a recipient of a New Investigator Award in Alzheimer's disease from the American Federation for Aging Research. CC is a recipient of a BrightFocus Foundation Alzheimer's Disease Research Grant (A2013359S). This study was funded by the Spanish Ministry of Science and Innovation SAF2006-10126 (2006–2009), SAF2010-22329-C02-01 (2010–2012), and SAF2013-47939-R (2013–2016) to PP. The list the full names of all of the PPMI funding partners found at www.ppmi-info.org/fundingpartners.

ACKNOWLEDGMENTS

The authors thank Susan Loftin and Karen Klumpp (Washington University School of Medicine) for their expert technical assistance. The authors also thank the Parkinson's Progression Markers Initiative (PPMI) project for granting access to data used in the preparation of this article. PPMI did not participate in the analyses or writing of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2018.00230/full#supplementary-material

REFERENCES

Aarsland, D., and Kurz, M. W. (2010). The epidemiology of dementia associated with Parkinson's disease. *Brain Pathol.* 20, 633–639. doi: 10.1111/j.1750-3639.2009.00369.x.

Aken, B. L., Ayling, S., Barrell, D., Clarke, L., Curwen, V., Fairley, S., et al. (2016). The Ensembl gene annotation system. *Database* 2016:baw093. doi: 10.1093/database/baw093

Amendola, L. M., Dorschner, M. O., Robertson, P. D., Salama, J. S., Hart, R., Shirts, B. H., et al. (2015). Actionable exomic incidental findings in 6503 participants: challenges of variant classification. *Genome Res.* 25, 305–315. doi: 10.1101/gr.183483.114

Baizabal-Carvallo, J. F., and Jankovic, J. (2016). Parkinsonism, movement disorders and genetics in frontotemporal dementia. *Nat. Rev. Neurol.* 12, 175–185. doi: 10.1038/nrneurol.2016.14

Benitez, B. A., and Cruchaga, C. (2013). TREM2 and neurodegenerative disease. N. Engl. J. Med. 369, 1567–1578 doi: 10.1056/NEJMc1306509#SA4

Benitez, B. A., Davis, A. A., Jin, S. C., Ibanez, L., Ortega-Cubero, S., Pastor, P., et al. (2016). Resequencing analysis of five Mendelian genes and the top genes from genome-wide association studies in Parkinson's disease. *Mol. Neurodegener.* 11:29. doi: 10.1186/s13024-016-0097-0

Bourdenx, M., Bezard, E., and Dehay, B. (2014). Lysosomes and Î±-synuclein form a dangerous duet leading to neuronal cell death. *Front. Neuroanat.* 8:83. doi: 10.3389/fnana.2014.00083

Brouwers, N., Nuytemans, K., van der Zee, J., Gijselinck, I., Engelborghs, S., Theuns, J., et al. (2007). Alzheimer and Parkinson diagnoses in progranulin null mutation carriers in an extended founder family. *Arch. Neurol.* 64, 1436–1446. doi: 10.1001/archneur.64.10.1436

Burdick, D. J., Cholerton, B., Watson, G. S., Siderowf, A., Trojanowski, J. Q., Weintraub, D., et al. (2014). People with Parkinson's disease and normal MMSE score have a broad range of cognitive performance. *Mov. Disord.* 29, 1258–1264. doi: 10.1002/mds.25924

Carecchio, M., Picillo, M., Valletta, L., Elia, A. E., Haack, T. B., Cozzolino, A., et al. (2017). Rare causes of early-onset dystonia-parkinsonism with cognitive impairment: a *de novo* PSEN-1 mutation. *Neurogenetics* 18, 175–178. doi: 10.1007/s10048-017-0518-4

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:7. doi: 10.1186/s13742-015-0047-8

Chang, D., Nalls, M. A., Hallgrímsdóttir, I. B., Hunkapiller, J., van der Brug, M., Cai, F., et al. (2017). A Meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nat. Genet.* 49, 1511–1516. doi: 10.1038/ng.3955

Collins, L. M., and Williams-Gray, C. H. (2016). The genetic basis of cognitive impairment and dementia in parkinson's disease. *Front. Psychiatry* 7:89. doi: 10.3389/fpsyt.2016.00089

Compta, Y., Parkkinen, L., O'Sullivan, S. S., Vandrovcova, J., Holton, J. L., Collins, C., et al. (2011). Lewy- and Alzheimer-type pathologies in Parkinson's disease dementia: Which is more important? *Brain* 134, 1493–1505. doi: 10.1093/brain/awr031

Correia Guedes, L., Ferreira, J. J., Rosa, M. M., Coelho, M., Bonifati, V., and Sampaio, C. (2010). Worldwide frequency of G2019S LRRK2 mutation in Parkinson's disease: A systematic review. *Park. Relat. Disord.* 16, 237–242. doi: 10.1016/j.parkreldis.2009.11.004

Cruchaga, C., Chakraverty, S., Mayo, K., Vallania, F. L., Mitra, R. D., Faber, K., et al. (2012a). Rare variants in APP, PSEN1 and PSEN2 increase risk for AD in late-onset Alzheimer's disease families. *PLoS ONE* 7:e31039. doi: 10.1371/journal.pone.0031039

Cruchaga, C., Kauwe, J. S., Nowotny, P., Bales, K., Pickering, E. H., Mayo, K., et al. (2012b). Cerebrospinal fluid APOE levels: An endophenotype for genetic studies for Alzheimer's disease. *Hum. Mol. Genet.* 21, 4558–4571. doi: 10.1093/hmg/dds296

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

Cuervo, A. M., Stefanis, L., Fredenburg, R., Lansbury, P. T., and Sulzer, D. (2004). Impaired degradation of mutant alpha-synuclein by chaperone-mediated autophagy. *Science* 305, 1292–1295. doi: 10.1126/science.1101738

Dalrymple-Alford, J. C., MacAskill, M. R., Nakas, C. T., Livingston, L., Graham, C., Crucian, G. P., et al. (2010). The MoCA: Well-suited screen for cognitive impairment in Parkinson disease. *Neurology* 75, 1717–1725. doi: 10.1212/WNL.0b013e3181fc29c9

Davis, A. A., Andruska, K. M., Benitez, B. A., Racette, B. A., Perlmutter, J. S., and Cruchaga, C. (2015). Variants in GBA, SNCA, and MAPT influence Parkinson disease risk, age at onset, and progression. *Neurobiol. Aging.* 37:209.e1–209.e7. doi: 10.1016/j.neurobiolaging.2015.09.014

Davydov, E. V., Goode, D. L., Sirota, M., Cooper, G. M., Sidow, A., and Batzoglou, S. (2010). Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput. Biol.* 6:e1001025. doi: 10.1371/journal.pcbi.1001025

Dehay, B., Martinez-Vicente, M., Caldwell, G. A., Caldwell, K. A., Yue, Z., Cookson, M. R., et al. (2013). Lysosomal impairment in Parkinson's disease. *Mov. Disord.* 28, 725–732. doi: 10.1002/mds.25462

Dubois, B., Burn, D., Goetz, C., Aarsland, D., Brown, R. G., Broe, G. A., et al. (2007). Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder society task force. *Mov. Disord.* 22, 2314–2324. doi: 10.1002/mds.21844

Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., Mizuno, Y., et al. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov. Disord.* 22, 1689–1707.

Federoff, M., Jimenez-Rolando, B., Nalls, M. A., and Singleton, A. B. (2012). A large study reveals no association between APOE and Parkinson's disease. *Neurobiol. Dis.* 46, 389–392. doi: 10.1016/j.nbd.2012.02.002

Fernández, M. V., Kim, J. H., Budde, J. P., Black, K., Medvedeva, A., Saef, B., et al. (2017). Analysis of neurodegenerative Mendelian genes in clinically diagnosed Alzheimer Disease. *PLoS Genet.* 13:e1007045. doi: 10.1371/journal.pgen.1007045

Ferrari, R., Wang, Y., Vandrovcova, J., Guelfi, S., Witeolar, A., Karch, C. M., et al. (2017). Genetic architecture of sporadic frontotemporal dementia and overlap with Alzheimer's and Parkinson's diseases. *J. Neurol. Neurosurg. Psychiatry* 88, 152–164. doi: 10.1136/jnnp-2016-314411

Foltynie, T., Brayne, C. E. G., Robbins, T. W., and Barker, R. A. (2004). The cognitive ability of an incident cohort of Parkinson's patients in the UK. The CamPaIGN study. *Brain* 127, 550–560. doi: 10.1093/brain/awh067

Goldman, J. G., and Litvan, I. (2011). Mild cognitive impairment in Parkinson's disease. *Minerva Med.* 102, 441–459.

Guerreiro, R., Escott-Price, V., Darwent, L., Parkkinen, L., Ansorge, O., Hernandez, D. G., et al. (2015). Genome-wide analysis of genetic correlation in dementia with Lewy bodies, Parkinson's and Alzheimer's diseases. *Neurobiol. Aging.* 38, 214.e7–214.e10. doi: 10.1016/j.neurobiolaging.2015.10.028

Harms, M. B., Neumann, D., Benitez, B. A., Cooper, B., Carrell, D., Racette, B. A., et al. (2013). Parkinson disease is not associated with C9ORF72 repeat expansions. *Neurobiol. Aging* 34:1519.e1-2. doi: 10.1016/j.neurobiolaging.2012.10.001

Healy, D. G., Falchi, M., O'Sullivan, S. S., Bonifati, V., Durr, A., Bressman, S., et al. (2008). Phenotype, genotype, and worldwide genetic penetrance of LRRK2-associated Parkinson's disease: a case-control study. *Lancet Neurol.* 7, 583–590. doi: 10.1016/S1474-4422(08)70117-0

Hoops, S., Nazem, S., Siderowf, A. D., Duda, J. E., Xie, S. X., Stern, M. B., et al. (2009). Validity of the MoCA and MMSE in the detection of MCI and dementia in Parkinson disease. *Neurology* 73, 1738–1745. doi: 10.1212/WNL.0b013e3181c34b47

Huang, X., Chen, P., Kaufer, D. I., Tröster, A. I., and Poole, C. (2006). Apolipoprotein E and dementia in Parkinson disease: a meta-analysis. *Arch. Neurol.* 63, 189–193. doi: 10.1001/archneur.63.2.189

Hughes, A. J., Daniel, S. E., Kilford, L., and Lees, A. J. (1992). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J. Neurol. Neurosurg. Psychiatry* 55, 181–184.

Irwin, D. J., Lee, V. M.-Y., and Trojanowski, J. Q. (2013). Parkinson's disease dementia: convergence of α-synuclein, tau and amyloid-β pathologies. *Nat. Rev. Neurosci.* 14, 626–636. doi: 10.1038/nrn3549

Jimenez-Escrig, A., Rabano, A., Guerrero, C., Simon, J., Barquero, M. S., Güell, I., et al. (2004). New V272A presenilin 1 mutation with very early onset subcortical dementia and parkinsonism. *Eur. J. Neurol.* 11, 663–669. doi: 10.1111/j.1468-1331.2004.00865.x

Jin, S. C., Carrasquillo, M. M., Benitez, B. A., Skorupa, T., Carrell, D., Patel, D., et al. (2015). TREM2 is associated with increased risk for Alzheimer's disease in African Americans. *Mol. Neurodegener.* 10, 19. doi: 10.1186/s13024-015-0016-9

Jin, S. C., Pastor, P., Cooper, B., Cervantes, S., Benitez, B. A., Razquin, C., et al. (2012). Pooled-DNA sequencing identifies novel causative variants in PSEN1, GRN and MAPT in a clinical early-onset and familial Alzheimer's disease Ibero-American cohort. *Alzheimer Res. Ther.* 4:34. doi: 10.1186/alzrt137

Josephs, K. A., Ahmed, Z., Katsuse, O., Parisi, J. F., Boeve, B. F., Knopman, D. S., et al. (2007). Neuropathologic features of frontotemporal lobar degeneration with ubiquitin-positive inclusions with progranulin gene (PGRN) mutations. *J. Neuropathol. Exp. Neurol.* 66, 142–151. doi: 10.1097/nen.0b013e31803020cf

Kao, A. W., McKay, A., Singh, P. P., Brunet, A., and Huang, E. J. (2017). Progranulin, lysosomal regulation and neurodegenerative disease. *Nat. Rev. Neurosci.* 18, 325–333. doi: 10.1038/nrn.2017.36

Kauwe, J. S. K., Jacquart, S., Chakraverty, S., Wang, J., Mayo, K., Fagan, A. M., et al. (2007). Extreme cerebrospinal fluid amyloid β levels identify family with late-onset Alzheimer's disease presenilin 1 mutation. *Ann. Neurol.* 61, 446–453. doi: 10.1002/ana.21099

Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., and 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

Kotzbauer, P. T., Cairns, N. J., Campbell, M. C., Willis, A. W., Racette, B. A., Tabbal, S. D., et al. (2012). Pathologic accumulation of α-synuclein and Aβ in Parkinson disease patients with dementia. *Arch. Neurol.* 69, 1326–1331. doi: 10.1001/archneurol.2012.1608

Kumar-Singh, S., Theuns, J., Van Broeck, B., Pirici, D., Vennekens, K., Corsmit, E., et al. (2006). Mean age-of-onset of familial alzheimer disease caused by presenilin mutations correlates with both increased Abeta42 and decreased Abeta40. *Hum Mutat* 27, 686–695. doi: 10.1002/humu.20336

Landrum, M. J., Lee, J. M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., et al. (2016). ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* 44, D862–D868. doi: 10.1093/nar/gkv1222

Lao, J. I., Beyer, K., Fernández-Novoa, L., and Cacabelos, R. (1998). A novel mutation in the predicted TM2 domain of the presenilin 2 gene in a Spanish patient with late-onset Alzheimer's disease. *Neurogenetics* 1, 293–296.

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291. doi: 10.1038/nature19057

Li, D., Parks, S. B., Kushner, J. D., Nauman, D., Burgess, D., Ludwigsen, S., et al. (2006). Mutations of presenilin genes in dilated cardiomyopathy and heart failure. *Am. J. Hum. Genet.* 79, 1030–1039. doi: 10.1086/509900

Liu, G., Boot, B., Locascio, J. J., Jansen, I. E., Winder-Rhodes, S., Eberly, S., et al. (2016). Specifically neuropathic Gaucher's mutations accelerate cognitive decline in Parkinson's. *Ann. Neurol.* 80, 674–685. doi: 10.1002/ana.24781

Lohmann, E., Guerreiro, R. J., Erginel-Unaltuna, N., Gurunlian, N., Bilgic, B., Gurvit, H., et al. (2012). Identification of PSEN1 and PSEN2 gene mutations and variants in Turkish dementia patients. *Neurobiol. Aging* 33:1850.e17–1850.e27. doi: 10.1016/j.neurobiolaging.2012.02.020

Mata, I. F., Johnson, C. O., Leverenz, J. B., Weintraub, D., Trojanowski, J. Q., Van Deerlin, V. M., et al. (2017). Large-scale exploratory genetic analysis of cognitive impairment in Parkinson's disease. *Neurobiol. Aging* 56, 211.e1–211.e7. doi: 10.1016/j.neurobiolaging.2017.04.009

Meeus, B., Theuns, J., and Van Broeckhoven, C. (2012a). The genetics of dementia with Lewy bodies: what are we missing? *Arch. Neurol.* 69, 1113–1118. doi: 10.1001/archneurol.2011.3678

Meeus, B., Verstraeten, A., Crosiers, D., Engelborghs, S., Van den Broeck, M., Mattheijssens, M., et al. (2012b). DLB and PDD: a role for mutations in dementia and Parkinson disease genes? *Neurobiol. Aging* 33:629.e5–629.e18. doi: 10.1016/j.neurobiolaging.2011.10.014

Morley, J. F., Xie, S. X., Hurtig, H. I., Stern, M. B., Colcher, A., Horn, S., et al. (2012). Genetic influences on cognitive decline in Parkinson's disease. *Mov. Disord.* 27, 512–518. doi: 10.1002/mds.24946

Nalls, M. A., Bras, J., Hernandez, D. G., Keller, M. F., Majounie, E., Renton, A. E., et al. (2015). NeuroX, a fast and efficient genotyping platform for investigation of neurodegenerative diseases. *Neurobiol. Aging* 36, 1605.e7–1605.e12. doi: 10.1016/j.neurobiolaging.2014.07.028

Nalls, M. A., Duran, R., Lopez, G., Kurzawa-Akanbi, M., McKeith, I. G., Chinnery, P. F., 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, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C., et al. (2009). Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 461, 272–276. doi: 10.1038/nature08250

Niwa, A., Matsuo, K., Shindo, A., Yata, K., Shiraishi, T., and Tomimoto, H. (2013). Clinical and neuropathological findings in a patient with familial Alzheimer disease showing a mutation in the PSEN1 gene. *Neuropathology* 33, 199–203. doi: 10.1111/j.1440-1789.2012.01340.x

Parsian, A., Racette, B., Goldsmith, L. J., and Perlmutter, J. S. (2002). Parkinson's disease and apolipoprotein E: possible association with dementia but not age at onset. *Genomics* 79, 458–461. doi: 10.1006/geno.2002.6707

Pastor, P., Ezquerra, M., Muñoz, E., Martí, M. J., Blesa, R., Tolosa, E., et al. (2000). Significant association between the tau gene A0/A0 genotype and Parkinson's disease. *Ann. Neurol.* 47, 242–245. doi: 10.1002/1531-8249(200002)47:2<242::AID-ANA16>3.0.CO;2-L

Petrucci, S., Consoli, F., and Valente, E. M. (2014). Parkinson Disease genetics: a "Continuum" from mendelian to multifactorial inheritance. *Curr. Mol. Med.* 14, 1079–1088. doi: 10.2174/1566524014666141010155509

Price, A., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N., a, and Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* 38, 904–909. doi: 10.1038/ng1847

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

Puschmann, A., Ross, O. A., Vilariño-Güell, C., Lincoln, S. J., Kachergus, J. M., Cobb, S. A., et al. (2009). A Swedish family with *de novo* α-synuclein A53T mutation: Evidence for early cortical dysfunction. *Park. Relat. Disord.* 15, 627–632. doi: 10.1016/j.parkreldis.2009.06.007

Qiao, Y., Peng, D., Jin, M., and Xue, S. (2017). Presenilin 1 Mutation (A431V) causing features of dementia with lewy bodies in a Chinese family of Alzheimer's Disease. *J. Alzheimer Dis. Park.* 7, 1–4. doi: 10.4172/2161-0460.1000307

Ratnavalli, E., Brayne, C., Dawson, K., and Hodges, J. R. (2002). The prevalence of frontotemporal dementia. *Neurology* 58, 1615–1621. doi: 10.1212/WNL.58.11.1615

Roberts, A. M., Ware, J. S., Herman, D. S., Schafer, S., Baksi, J., Bick, A. G., et al. (2015). Integrated allelic, transcriptional, and phenomic dissection of the cardiac effects of titin truncations in health and disease. *Sci. Transl. Med.* 7:270ra6. doi: 10.1126/scitranslmed.3010134

Rocca, W. A., Bower, J. H., Ahlskog, J. E., Elbaz, A., Grossardt, B. R., McDonnell, S. K., et al. (2007). Risk of cognitive impairment or dementia in relatives of patients with Parkinson disease. *Arch. Neurol.* 64, 1458–1464. doi: 10.1001/archneur.64.10.1458

Rovelet-Lecrux, A., Deramecourt, V., Legallic, S., Maurage, C. A., Le Ber, I., Brice, A., et al. (2008). Deletion of the progranulin gene in patients with frontotemporal lobar degeneration or Parkinson disease. *Neurobiol. Dis.* 31, 41–45. doi: 10.1016/j.nbd.2008.03.004

Sannerud, R., Esselens, C., Ejsmont, P., Mattera, R., Rochin, L., Tharkeshwar, A. K., et al. (2016). Restricted location of PSEN2/γ-secretase determines substrate specificity and generates an intracellular Aβ pool. *Cell* 166, 193–208. doi: 10.1016/j.cell.2016.05.020

Schulte, E. C., Fukumori, A., Mollenhauer, B., Hor, H., Arzberger, T., Perneczky, R., et al. (2015). Rare variants in β-Amyloid precursor protein (APP) and Parkinson's disease. *Eur. J. Hum. Genet.* 23, 1328–1333. doi: 10.1038/ejhg.2014.300

Setó-Salvia, N., Pagonabarraga, J., Houlden, H., Pascual-Sedano, B., Dols-Icardo, O., Tucci, A., et al. (2012). Glucocerebrosidase mutations confer a greater risk of dementia during Parkinson's disease course. *Mov. Disord.* 27, 393–399. doi: 10.1002/mds.24045

Snider, B. J., Norton, J., Coats, M., a, Chakraverty, S., Hou, C. E., Jervis, R., et al. (2005). Novel presenilin 1 mutation (S170F) causing Alzheimer disease with lewy bodies in the third decade of life. *Arch. Neurol.* 62, 1821–1830. doi: 10.1001/archneur.62.12.1821

Srivatsal, S., Cholerton, B., Leverenz, J. B., Wszolek, Z. K., Uitti, R. J., Dickson, D. W., et al. (2015). Cognitive profile of LRRK2-related Parkinson's disease. *Mov. Disord.* 30, 728–733. doi: 10.1002/mds.26161

Takao, M., Ghetti, B., Hayakawa, I., Ikeda, E., Fukuuchi, Y., Miravalle, L., et al. (2002). A novel mutation (G217D) in the Presenilin 1 gene (PSEN1) in a Japanese family: Presenile dementia and parkinsonism are associated with cotton wool plaques in the cortex and striatum. *Acta Neuropathol.* 104, 155–170. doi: 10.1007/s00401-002-0536-6

Tanaka, Y., Matsuwaki, T., Yamanouchi, K., and Nishihara, M. (2013). Increased lysosomal biogenesis in activated microglia and exacerbated neuronal damage after traumatic brain injury in progranulin-deficient mice. *Neuroscience* 250, 8–19. doi: 10.1016/j.neuroscience.2013.06.049

Terrelonge, M., Marder, K. S., Weintraub, D., and Alcalay, R. N. (2015). CSF β-amyloid 1-42 predicts progression to cognitive impairment in newly diagnosed Parkinson disease. *J. Mol. Neurosci.* 58, 88–92. doi: 10.1007/s12031-015-0647-x

Thaler, A., Mirelman, A., Gurevich, T., Simon, E., Orr-Urtreger, A., Marder, K., et al. (2012). Lower cognitive performance in healthy G2019S LRRK2 mutation carriers. *Neurology* 79, 1027–1032. doi: 10.1212/WNL.0b013e3182684646

Tomaino, C., Bernardi, L., Anfossi, M., Costanzo, A., Ferrise, F., Gallo, M., et al. (2007). Presenilin 2 Ser130Leu mutation in a case of late-onset "sporadic" Alzheimer's disease [3]. *J. Neurol.* 254, 391–393. doi: 10.1007/s00415-006-0373-y

Tsuang, D., Leverenz, J. B., Lopez, O. L., Hamilton, R. L., Bennett, D. A., Schneider, J. A., et al. (2013). APOE epsilon4 increases risk for dementia in pure synucleinopathies. *JAMA Neurol.* 70, 223–228. doi: 10.1001/jamaneurol.2013.600

Vallania, F. L. M., Druley, T. E., Ramos, E., Wang, J., Borecki, I., Province, M., et al. (2010). High-throughput discovery of rare insertions and deletions in large cohorts. *Genome Res.* 20, 1711–1718. doi: 10.1101/gr.109157.110

Van Kampen, J. M., Baranowski, D., and Kay, D. G. (2014). Progranulin gene delivery protects dopaminergic neurons in a mouse model of Parkinson's disease. *PLoS ONE* 9:e97032. doi: 10.1371/journal.pone.0097032

Walker, E. S., Martinez, M., Brunkan, A. L., and Goate, A. (2005). Presenilin 2 familial Alzheimer's disease mutations result in partial loss of function and dramatic changes in Aβ 42/40 ratios. *J. Neurochem.* 92, 294–301. doi: 10.1111/j.1471-4159.2004.02858.x

Walsh, R., Thomson, K. L., Ware, J. S., Funke, B. H., Woodley, J., McGuire, K. J., et al. (2017). Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and 60,706 reference samples. *Genet. Med.* 19, 192–203. doi: 10.1038/gim.2016.90

Weintraub, D., Papay, K., and Siderowf, A. (2013). Screening for impulse control symptoms in patients with *de novo* Parkinson disease: a casecontrol study. *Neurology* 80, 176–180. doi: 10.1212/WNL.0b013e31827b915c

Winder-Rhodes, S. E., Evans, J. R., Ban, M., Mason, S. L., Williams-Gray, C. H., Foltynie, T., et al. (2013). Glucocerebrosidase mutations influence the natural history of Parkinson's disease in a community-based incident cohort. *Brain* 136, 392–399.

Wright Willis, A., Evanoff, B. A., Lian, M., Criswell, S. R., and Racette, B. A. (2010). Geographic and ethnic variation in Parkinson disease: A population-based study of us medicare beneficiaries. *Neuroepidemiology* 34, 143–151. doi: 10.1159/000275491

Wu, M. C., Lee, S., Cai, T., Li, Y., Boehnke, M., and Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. *Am. J. Hum. Genet.* 89, 82–93. doi: 10.1016/j.ajhg.2011.05.029

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 30 April 2018
doi: 10.3389/fnins.2018.00260

Combining Polygenic Hazard Score With Volumetric MRI and Cognitive Measures Improves Prediction of Progression From Mild Cognitive Impairment to Alzheimer's Disease

Karolina Kauppi[1,2], Chun Chieh Fan[1,3], Linda K. McEvoy[1], Dominic Holland[4], Chin Hong Tan[5], Chi-Hua Chen[1], Ole A. Andreassen[6,7], Rahul S. Desikan[5], and Anders M. Dale[1,3,4]* for the Alzheimer's Disease Neuroimaging Initiative[†]*

[1] Department of Radiology, University of California, San Diego, La Jolla, CA, United States, [2] Department of Radiation Sciences, University of Umea, Umea, Sweden, [3] Department of Cognitive Sciences, University of California, San Diego, La Jolla, CA, United States, [4] Department of Neurosciences, University of California, San Diego, La Jolla, CA, United States, [5] Neuroradiology Section, Department of Radiology and Biomedical Imaging, University of California, San Francisco, San Francisco, CA, United States, [6] NORMENT, Institute of Clinical Medicine, Division of Mental Health and Addiction, University of Oslo, Oslo University Hospital, Oslo, Norway, [7] Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway

Improved prediction of progression to Alzheimer's Disease (AD) among older individuals with mild cognitive impairment (MCI) is of high clinical and societal importance. We recently developed a polygenic hazard score (PHS) that predicted age of AD onset above and beyond *APOE*. Here, we used data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) to further explore the potential clinical utility of PHS for predicting AD development in older adults with MCI. We examined the predictive value of PHS alone and in combination with baseline structural magnetic resonance imaging (MRI) data on performance on the Mini-Mental State Exam (MMSE). In survival analyses, PHS significantly predicted time to progression from MCI to AD over 120 months (p = 1.07e-5), and PHS was significantly more predictive than *APOE* alone (p = 0.015). Combining PHS with baseline brain atrophy score and/or MMSE score significantly improved prediction compared to models without PHS (three-factor model p = 4.28e-17). Prediction model accuracies, sensitivities and area under the curve were also improved by including PHS in the model, compared to only using atrophy score and MMSE. Further, using linear mixed-effect modeling, PHS improved the prediction of change in the Clinical Dementia Rating—Sum of Boxes (CDR-SB) score and MMSE over 36 months in patients with MCI at baseline, beyond both *APOE* and baseline levels of brain atrophy. These results illustrate the potential clinical utility of PHS for assessment of risk for AD progression among individuals with MCI both alone, or in conjunction with clinical measures of prodromal disease including measures of cognitive function and regional brain atrophy.

Keywords: pHs, MCI, AD prediction, MRI, genetics

OPEN ACCESS

Edited by:
Stefano L. Sensi,
Università degli Studi G. d'Annunzio Chieti e Pescara, Italy

Reviewed by:
Stefano Delli Pizzi,
Università degli Studi G. d'Annunzio Chieti e Pescara, Italy
Fabrizio Piras,
Fondazione Santa Lucia (IRCCS), Italy

****Correspondence:***
Karolina Kauppi
karolina.kauppi@umu.se
Anders M. Dale
amdale@ucsd.edu

[†]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Specialty section:
This article was submitted to Neurodegeneration, a section of the journal Frontiers in Neuroscience

Received: *09 January 2018*
Accepted: *04 April 2018*
Published: *30 April 2018*

Citation:
Kauppi K, Fan CC, McEvoy LK, Holland D, Tan CH, Chen C-H, Andreassen OA, Desikan RS, Dale AM and for the Alzheimer's Disease Neuroimaging Initiative (2018) Combining Polygenic Hazard Score With Volumetric MRI and Cognitive Measures Improves Prediction of Progression From Mild Cognitive Impairment to Alzheimer's Disease. Front. Neurosci. 12:260. doi: 10.3389/fnins.2018.00260

INTRODUCTION

Late onset Alzheimer's disease (AD) is the most common form of dementia, affecting 24–35 million people world-wide (Querfurth and LaFerla, 2010; Alzheimer's Association, 2015). Novel methods to enable early AD detection based on clinically feasible, economical, and non-invasive measures is of high clinical and societal value. Early identification of high-risk individuals is also of utmost importance for pre-dementia clinical trials (Holland et al., 2012; Ritchie et al., 2016).

A large effort has been made to improve the prediction of progression to AD among older individuals with mild cognitive impairment (MCI), which can be a transition stage from normal age-related cognitive decline to dementia (Roberts and Knopman, 2013). Next to older age, inheritance of the *ε4* allele of the *Apolipoprotein E* (*APOE)* gene is the strongest individual risk factor for late onset AD (Yu et al., 2014). However, other genetic risk variants of smaller effect also contribute to AD risk (Lambert et al., 2013; Ridge et al., 2013). While most studies examining genetic risk for sporadic AD focus on *APOE* genotype, some have assessed polygenic risk scores beyond *APOE* based on case-control data from large-scale genome-wide association studies (GWAS) (Lambert et al., 2013), and found that genetic variants associated with elevated AD risk also influence brain structure (Sabuncu et al., 2012) and cognitive function (Marioni et al., 2017). Given that the incidence of AD increases sharply with age, we recently develop a polygenic hazard score (PHS) for prediction of age-specific AD risk, based on 31 AD-susceptibility variants, including *APOE* (Desikan et al., 2017). The PHS showed substantial improvement over *APOE* in predicting age of AD onset and was associated with biomarkers of AD, including MRI-based hippocampal volume loss (Desikan et al., 2017), amyloid, and tau deposition (Tan et al., 2018).

Volumetric MRI-based measures of regional brain atrophy, particularly medial temporal volume loss are important biomarkers for assessing risk of progression to AD in patients with MCI (Li et al., 2016). In previous work from our lab, we derived a composite regional "brain atrophy score" from linear discrimination analysis trained on data from healthy controls and AD patients, which was better at predicting 1 year cognitive decline than atrophy in medial temporal structures alone (McEvoy et al., 2009). The atrophy score is based on volume of the hippocampus, and thickness of entorhinal cortex, middle temporal gyrus, bank of the superior temporal sulcus, isthmus cingulate (retrosplenial cortex), superior temporal gyrus, medial and lateral orbitofrontal gyri, with weightings for each ROI determined through a linear discrimination analysis that best distinguished AD patients from healthy controls. In a subsequent study, this brain atrophy score predicted 1-year risk of progression to AD in individual patients (McEvoy et al., 2011). Prediction of 1-year clinical decline was further improved by adding subjects' baseline Mini-Mental State Exam (MMSE) scores and number of *APOE ε4* alleles to the model (McEvoy et al., 2009).

In the current study, we investigated the clinical utility of PHS for individual assessment of risk for clinical progression to AD over time among older individuals with MCI, a critical question for most patients admitted to memory clinics. Whereas prior work from our group has shown the value of PHS for predicting AD-associated clinical and cognitive decline among non-demented elderly individuals (Tan et al., 2017), a critical next step in assessing the potential clinical utility of PHS is to examine the extent to which PHS provides independent information beyond other commonly used predictors, such as brain atrophy levels and baseline cognitive function, and to determine whether combinations of these measures improve prediction of clinical decline and progression to dementia. To this end, we used survival analyses (Klein et al., 2013) to compare single-, two-, and three-factor models of PHS, atrophy score, and MMSE for prediction of time to progression from MCI to AD over 120 months of follow-up. As a complementary approach, we used linear mixed-effect modeling to examine prediction of clinical change in MMSE and the Clinical Dementia Rating, sum of boxes (CDR-SB), over 36 months.

METHODS

Participants

We used participants from the ADNI database, available as of November 2011. ADNI 1 is a 5-year multi-site program launched in 2003 as a public-private partnership including the National Institute on Aging, Food and Drug Administration, pharmaceutical companies, and nonprofit organizations (led by Principal Investigator Michael W. Weiner, MD.). The main goal of ADNI is to examine if progression from MCI to AD can be predicted based on neuroimaging, biological biomarkers as well as clinical and neuropsychological assessments.

The baseline data was collected in 2005, including elderly healthy controls ($n = 200$), Alzheimer's disease patients ($n = 200$), and individuals with MCI ($n = 400$), followed by annual follow-ups for 36 months. We included 336 participant with a MCI diagnosis at baseline in ADNI 1 and available genetic, MRI, and cognitive data, including mini-mental state examination (MMSE) (Folstein et al., 1975) and Clinical dementia rating, sum of boxes (CDR-SB) (Hughes et al., 1982). The age range was 55–89 at baseline. Longitudinal data on CDR-SB and MMSE was included from ADNI 1, with 36-month follow-up. Data on progression to AD was also included from ADNI 2 and ADNI GO, providing data on progression to AD for up to 120 months after baseline. The ADNI study was approved by local institutional review boards, and all participants or participant's guardians provided written informed consent. Additional information about ADNI is available at http://www.adni-info.org.

MRI Acquisition and Analyses

Details of image acquisition and analysis have been described in our previous publications (McEvoy et al., 2009). Briefly, we downloaded the raw baseline DICOM MRI data from the ADNI web site (http://adni.loni.usc.edu/data-samples/mri/) and obtained volumetric assessments on neuroanatomic regions of interest (ROIs) using a modified version of the FreeSurfer image-analysis software (Brewer, 2009; Brewer et al., 2009). We used a previously validated brain atrophy score based on volume of

the hippocampus, and thickness of entorhinal cortex, middle temporal gyrus, bank of the superior temporal sulcus, isthmus cingulate (retrosplenial cortex), superior temporal gyrus, medial and lateral orbitofrontal gyri, with weightings for each ROI determined through a linear discrimination analysis that best distinguished AD patients from healthy controls (for additional details, see McEvoy et al., 2009, 2011). In the current paper, we computed an atrophy score as the sum of weighted measures from these brain regions, averaging left, and right hemispheres.

Polygenic Hazard Score (PHS)

For each participant included in the study, we calculated their individual PHS based on our previous publication (Desikan et al., 2017). In brief, PHS was derived by first identifying common variants associated with AD in the International genetics of Alzheimer's Project (IGAP) GWAS on AD, with a *p*-value threshold of $<10^{-5}$ using summary statistics. These SNPs were examined for association with AD in the Alzheimer's Disease Genetic consortia (ADGC) phase 1 genetic data. A stepwise Cox proportional hazard model was applied in which the SNPs that improved the model most were included sequentially until the model was no longer improved by adding more SNPs. This resulted in a list of 31 SNPs, including the two SNPs that constitute the *APOE* ε genotype, which were used to generate the PHS. Finally, the PHS estimates from ADGC were integrated with established AD-incidence rates from the US population to provide quantitative estimates of the annualized (cumulative) incidence rate. The PHS is the vector product of a person's genotype for the 31 SNPs and the corresponding parameter estimates from the Cox proportional hazard model.

Statistical Analyses

We used Cox proportional Hazard models in Matlab [version 8.5.0.197613 (R2015a)] to model time to progression from MCI to AD with a follow up period of 120 months. Time to event was defined as time from baseline to AD onset (with end of study time or drop out as censoring). We used Kaplan–Meier survival analysis to determine the time to progression for the 10th, 50th, and 90th risk percentile. We first fitted a baseline model containing age, age^2, sex, and age*sex as predictors. We then added PHS to the baseline model, and used log likelihood ratio to assess whether PHS significantly improved the prediction. Model comparisons between PHS and number of *APOE* ε4 alleles were also made (Baseline and *APOE* vs. Baseline, *APOE* and PHS). Thereafter, we examined the combined model of PHS and atrophy score at baseline, and performed model comparisons with each factor alone. Finally, we added baseline measure of MMSE to a three-factor model, and performed model comparisons among all two-factor models. Cross-validated prediction accuracy, sensitivity, specificity and area under the curve (AUC) were calculated via the receiver operator characteristics (ROC) analyses using the perfcurve function in Matlab (MathWorks). The assumption of proportional hazards was not violated for any of the included covariates (p's > 0.05, as evaluated by scaled Schoenfeld residuals using the cox.zph function in R).

To predict cognitive decline in patients with MCI, we used linear mixed-effects models to estimate change in MMSE or CDR-SB over 36 months. Mixed effects models were fitted via maximum likelihood by using the lmer function in R (version 3.2.3). Sex, age, education, and five genetic principal components to control for population stratification were included in all analyses as baseline variables. We included PHS, atrophy score and the combination of PHS and atrophy score in three different models each for prediction of change in CDR-SB and MMSE. The models allowed for random subject-specific intercept and slope. Model comparisons of *APOE* and *APOE*+PHS were performed, where *APOE* denotes the number of *APOE* ε4 alleles. Likelihood ratio test via ANOVA were used for all model comparisons.

RESULTS

Of the 336 participants with MCI at baseline, 182 developed AD within the follow-up period of 120 months. Baseline demographics of stable MCI (MCI-s) and MCI patients that subsequently progressed to AD (MCI-c) are presented in **Table 1**. The groups did not differ in sex distribution, age, or education. As expected, *APOE* ε4 alleles, MMSE, atrophy score, and PHS were related to subsequent progression.

Estimated survival functions for the 10th, 50th, and 90th percentile based on one-, two-, and three-factors cox proportional hazard models modeling time to progression from MCI to AD are shown in **Figure 1** (**Figure 1A**, PHS, **Figure 1B**, PHS and atrophy score, **Figure 1C**, PHS, atrophy score, and MMSE), and model summaries are shown in **Table 2**. The PHS significantly predicted progression from MCI to AD over 120 months follow-up (p = 1.07e-5), and PHS was a significantly stronger predictor of progression than *APOE* ε genotype ($p = 0.0152$, for model comparison of *APOE* vs. *APOE* +PHS). When including atrophy score (McEvoy et al., 2009) in the model, PHS remained significant and the two-factor prediction model was significantly more predictive than either single-factor model (p's = 5.61e-11, and 0.0015 for comparison with

TABLE 1 | Clinical demographics.

	MCIs (n = 154)	MCIc (n = 182)	Statistics	p
Males, *n* (%)	99 (64%)	117 (64%)	$\chi^2_{(1)} = 0$	1
Age, y (*SD*)	75.84 (7.4)	74.92 (6.9)	$t_{(334)} = -1.18$	0.24
APOE4 +, *n* (% E4+)	69 (45%)	117 (64%)	$\chi^2_{(1)} = 12$,	$5.2e^{-4}$
Education	15.68	15.82	$t_{(334)} = 0.41$	0.68
MMSE	27.37	26.87	$t_{(333)} = -2.59$	0.01
Atrophy score	2382	2207	$t_{(330)} = -4.45$	$1.17e^{-5}$
PHS	0.356	0.661	$t_{(334)} = 3.54$	$4.42e^{-4}$

Clinical and demographics data at baseline for patients with stable MCI (MCIs) and those who converted to AD within the study period (MCIc). MCI, mild cognitive impairment; PHS, Polygenic hazard score; SD, standard deviation.

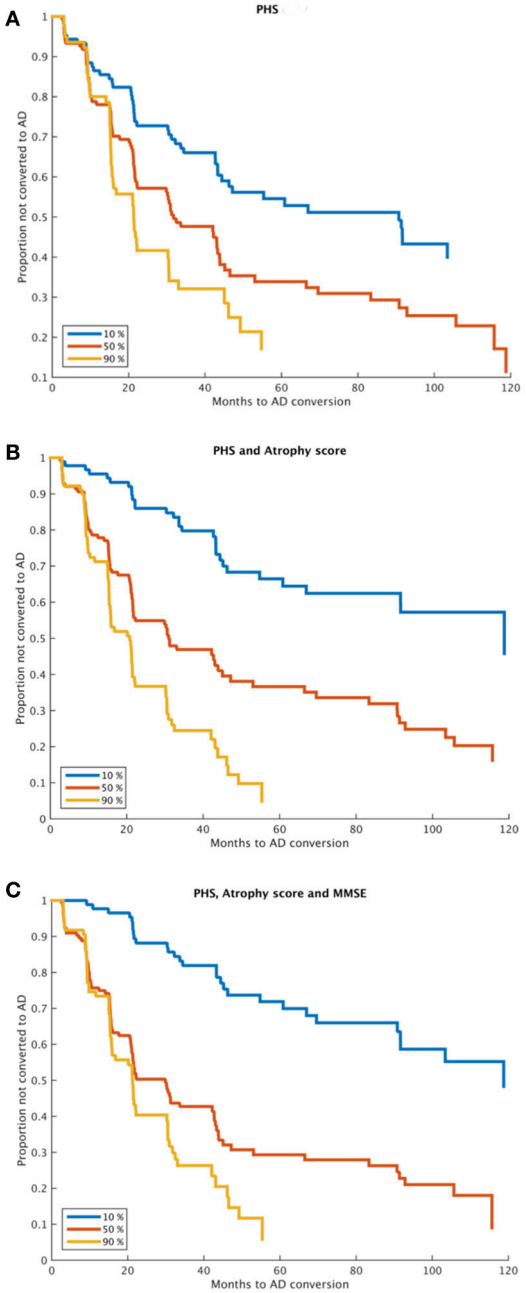

FIGURE 1 | Time to MCI to AD progression. Kaplan-Meier survival curves for 10, 50, and 90th percentiles of **(A)** PHS, **(B)** PHS and atrophy score, and **(C)** PHS, atrophy score and MMSE (at baseline). Model comparisons using log likelihood of the two-factor model **(B)** vs. any one-factor model, and the three-factor model **(C)** vs. any combination of two-factor models; all $p <$ 0.005. PHS, Polygenic Hazard Score; AD, Alzheimer's Disease.

single-factor models of PHS and atrophy score, respectively). Finally, we included cognitive functioning at baseline (MMSE) to a three-factor prediction model, which yielded a combined model p-value of 4.28e-17. Model comparisons showed that the three-factor model was significantly more predictive than the two-factor model ($p < 0.005$).

ROC analyses were performed to assess performance accuracies for one-, two- and three-factor models for prediction of progression over a 36 months follow-up period based on combinations of atrophy score, MMSE, and PHS (**Table 3**). In brief, the atrophy score had an accuracy of 74.6, sensitivity of 77.8, and a specificity of 70.8. Adding PHS to the atrophy score increased specificity, at the cost of sensitivity; with increased overall accuracy and AUC (although confidence intervals overlapped). Similar results were seen for comparisons of MMSE alone and in combination with PHS. The full three-factor model had the highest AUC (.84), accuracy (78.9), sensitivity (79.9), and a specificity of 77.8.

Results from the linear mixed-effect models for prediction of change in MMSE and CDR-SB over 36 months are presented in **Table 4**. Model comparisons showed that PHS significantly improved prediction of both MMSE ($\chi^2 = 26.7$, df $= 1$, $p = 2.34e\text{-}07$) and CDR-SB ($\chi^2 = 21.57$, df $= 1$, $p = 3.41e\text{-}06$) compared to the baseline variables. Further, the PHS performed significantly better than *APOE* ε4 status in prediction of both MMSE ($\chi^2 = 8.61$, df $= 1$, $p = 0.0033$) and CDR-SB ($\chi^2 = 6.12$, df $= 1$, $p = 0.013$). Again, PHS remained significant after adding atrophy score to the model (**Table 4**). Compared to atrophy score alone, the combined model of PHS and atrophy score was significantly more predictive of change in both MMSE ($\chi^2 = 19.04$, df $= 1$, $p = 1.281e\text{-}05$, [controlling for *APOE* ε4 alleles: $\chi^2 = 6.97$, df $= 1$, p $= 0.008$]) as well as CDR_SB ($\chi^2 = 13.43$, df $= 1$, $p = 0.00025$ [controlling for APOE ε4 alleles: $\chi^2 = 4.57$, df $= 1$, $p = 0.033$]).

DISCUSSION

Using the ADNI dataset, we assessed the potential clinical utility of the recently established age-specific AD PHS (Desikan et al., 2017) to more accurately predict the progression to AD among patients diagnosed with MCI. Examining both time to AD diagnosis and decline in clinical scores (CDR-SB and MMSE), we found that PHS significantly predicted clinical progression to AD beyond *APOE* genotype. Critically, PHS remained a significant predictor even when regional brain atrophy levels and cognitive functioning was known, and the prediction models were significantly improved by adding PHS to models containing atrophy score and/or MMSE. These results show that the predictive value of PHS is, at least partly, independent of brain atrophy and cognitive functioning at the MCI stage. In a typical memory clinic setting, PHS may be used to assess AD risk before other diagnostics have been performed, and may also be valuable for further improvement of prediction after an individual has underwent an MRI examination.

We used two complementary approaches to assess clinical progression among MCI patients. First, we used survival analyses to assess time to progression from MCI to AD over a long period of 120 months, using combinations of PHS, MRI, and cognitive functioning (**Figure 1**). The PHS was a significant predictor both alone and when controlling for atrophy level and cognitive function at study start, and model comparisons showed significant model improvements

TABLE 2 | Results from Cox model analyses.

Variables	Log(HR)	SE	*t*	*p*
BASELINE VARIABLES				
Age	−0.229	0.192	−1.190	0.234
Age^2	0.001	0.001	0.907	0.364
Sex	−2.788	1.706	−1.634	0.102
Sex*age	0.035	0.022	1.554	0.120
SINGLE-FACTOR MODELS				
PHS	0.454	0.097	4.684	2.81E-06
Atrophy score	−0.002	2.45E-04	−7.276	3.44E-13
MMSE	−0.205	0.044	−4.629	3.68E-06
TWO-FACTOR MODEL				
PHS	0.314	0.099	3.175	0.001
Atrophy score	−0.002	2.54E-04	−6.463	1.03E-10
THREE-FACTOR MODEL				
PHS	0.264	0.100	2.644	0.008
Atrophy score	−0.002	2.58E-04	−5.962	2.49E-09
MMSE	−0.142	0.047	−3.010	0.003

Baseline variables incorporated in all subsequent models. PHS, Polygenic hazard score; MMSE, mini-mental state examination; HR, Hazard ratio.

TABLE 3 | Prediction accuracy.

Variables	Accuracy (%)	Sensitivity (%)	Specificity (%)	AUC (95% CI)
PREDICTORS				
MMSE	68.4	62.8	75.1	0.73 (0.68–0.78)
Atrophy score	74.6	77.8	70.8	0.79 (0.74–0.83)
PHS and MMSE	70.0	57.2	85.1	0.79 (0.74–0.83)
PHS and atrophy score	76.1	72.9	79.8	0.82 (0.77–0.86)
MMSE and Atrophy score	77.2	77.08	77.3	0.82 (0.77–0.86)
Atrophy score, PHS and MMSE	78.9	79.9	77.8	0.84 (0.79–0.88)

Prediction of progression from MCI to AD by 36 months follow up. PHS, polygenic hazard score; MMSE, mini-mental state examination; MCI, mild cognitive impairment; AD, Alzheimer's Disease; AUC, area under the curve.

when adding PHS to models consisting of atrophy score and/or MMSE. ROC analyses showed that adding PHS to atrophy score or MMSE improved prediction accuracy, which was primarily driven by increased specificity, but at the cost of lower sensitivity (**Table 3**). These results suggest that different combinations of biomarkers may be used in different clinical situations where higher sensitivity or specificity is prioritized. As expected, the best model performance was derived from the full three- factor model with PHS, MRI, and cognitive assessment.

Secondly, we used linear mixed-effect models to examine the influence of PHS on change in CDR-SB and MMSE over a shorter time period, 3 years, in the same individuals with MCI at baseline. In line with the results from survival analyses, PHS significantly predicted clinical decline of both CDR-SB and MMSE among elderly diagnosed with MCI beyond *APOE* status,

TABLE 4 | Prediction of clinical decline.

Model	Predictors	Estimate	SE	*t*	*p*
OUTCOME: MMSE					
PHS	PHS	−0.87	0.17	−5.2(253.2)	4.06e−07
PHS+ Atrophy score	PHS	−0.72	0.17	−4.34 (251.5)	2.05e−05
	Atrophy score	2.51	0.67	3.73 (243.6)	0.000238
OUTCOME: CDR-SB					
PHS	PHS	0.46	0.10	4.65(259.3)	5.28e−06
PHS+ Atrophy score	PHS	0.42	0.10	4.27 (256.1)	2.80e−05
	Atrophy score	−1.79	0.38	4.70 (249.3)	4.39e-06

Effect of predictors on change in outcome measure over 36 months. PHS, Polygenic Hazard score; MMSE, mini mental state examination; CDR-SB, Clinical Dementia Rating score, sub of Boxes.

both individually and when atrophy levels were included in the model (**Table 4**). Taken together, these results show the utility of the PHS to assess individual risk for clinical decline in patients diagnosed with MCI also when their current levels of brain atrophy and cognitive functioning is known.

Previous studies using both survival analyses and linear mixed effect models showed improved prediction of AD progression by combining MRI data with CSF biomarkers (Vemuri et al., 2009; Westman et al., 2012), *APOE* genotype (McEvoy et al., 2009; Dukart et al., 2015), and different cognitive test batteries (Callahan et al., 2015; Eckerström et al., 2015; Li et al., 2016). Investigators from our group have shown that the combination of cognitive performance and medial temporal atrophy substantially improves prediction of MCI to AD progression in comparison to prediction based on individual risk factors (Heister et al., 2011), and also that the combination of *APOE* genotype and brain atrophy outperforms models based on either variable alone (McEvoy et al., 2009). Here, we extend previous models based on *APOE* to our recently developed PHS based on whole-genome data. It is still not fully known through which mechanisms *APOE* and other genes with smaller effect impact AD risk, but brain atrophy level and cognitive function are considered intermediate phenotypes that may mediate genetic effects on AD risk. In our previous paper, we found a correlation between PHS and larger volume loss in AD-related brain areas (Desikan et al., 2017). The current findings that PHS is predictive of AD progression when levels of brain atrophy are included in the model, shows that MRI biomarkers are not fully mediating the effect of PHS. Prediction based on genetic testing has the advantages of being relatively cheap, non-invasive, and not time-sensitive (since genetic assessment only has to be carried out once, is valid for a whole life time and can be used for multiple clinical purposes). In contrast to polygenic risk scores developed in a case/control framework, PHS is focused on predicting age of onset, which more accurately captures the increase in population incidence with increased age, where older age is the strongest risk factor for AD development.

LIMITATIONS

ADNI is one of the largest longitudinal dataset for studying progression from MCI to AD, but has limitations. First, participants were recruited from memory clinics and advertisements, and MCI inclusion criteria were highly selective, thus the study group is not representative of the general population. Further, AD diagnosis has not been confirmed with histopathology. Also, study dropouts are biased toward high-risk individuals, which might lead to a bias in the estimates of progression rates. As high-risk individuals are more likely to have a higher PHS, the predictive value of the PHS might be underestimated in this study.

CONCLUSIONS

The present study shows that the prediction of clinical progression to AD among MCI patients can be improved by combining the age-sensitive PHS with structural neuroimaging and baseline cognitive ability. Improved individual assessment of AD risk among elderly patients presenting with subjective memory complaints could be helpful in clinical practice to determine treatment plans, and is also of high importance for intervention studies where recruitment of high-risk individuals at an early stage of the disease process is crucial for testing effectiveness of new disease-altering interventions.

ETHICS STATEMENT

The ADNI study was approved by the local ethics committee at each center. All subjects gave written informed consent in accordance with the Declaration of Helsinki. For more information see ADNI webpage; http://adni.loni.usc.edu.

AUTHOR CONTRIBUTIONS

KK, CF, RD, and AD: designed the study; CF and KK: performed analyses; KK: wrote the manuscript; CF, LM, DH, CT, C-HC, OA, RD, and AD: read and provided comments on the manuscript.

ACKNOWLEDGMENTS

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. This research was also supported by NIH grants P30 AG010129 and K01 AG030514. KK was supported by a grant from the Swedish Research Council, OA by the Research Council of Norway and KG Jebsen Foundation, RD by the National Alzheimer's Coordinating Center Junior Investigator Award, RSNA Resident/Fellow Award Foundation and ASNR Alzheimer's Imaging Grant, and LM by the National Institute on Alcohol Abuse and Alcoholism (R01AA021187).

REFERENCES

Alzheimer's Association, A. (2015). 2015 Alzheimer' s disease facts and figures. *Alzheimer's Dement.* 11, 332–384. doi: 10.1016/j.jalz.2015.02.003

Brewer, J. B. (2009). Fully-automated volumetric MRI with normative ranges: translation to clinical practice. *Behav. Neurol.* 21, 21–28. doi: 10.3233/BEN-2009-0226

Brewer, J. B., Magda, S., Airriess, C., and Smith, M. E. (2009). Fully-automated quantification of regional brain volumes for improved detection of focal atrophy in Alzheimer disease. *AJNR Am. J. Neuroradiol.* 30, 578–580. doi: 10.3174/ajnr.A1402

Callahan, B. L., Ramirez, J., Berezuk, C., Duchesne, S., and Black, S. E. (2015). Predicting Alzheimer's disease development: a comparison of cognitive criteria and associated neuroimaging biomarkers. *Alzheimers Res. Ther.* 7:68. doi: 10.1186/s13195-015-0152-z

Desikan, R. S., Fan, C. C., Wang, Y., Schork, A. J., Cabral, H. J., Cupples, L. A., 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

Dukart, J., Sambataro, F., and Bertolino, A. (2015). Accurate prediction of conversion to Alzheimer's disease using imaging, genetic, and neuropsychological biomarkers. *J. Alzheimers Dis.* 49, 1143–1159. doi: 10.3233/JAD-150570

Eckerström, C., Olsson, E., Klasson, N., Berge, J., Nordlund, A., Bjerke, M., et al. (2015). Multimodal prediction of dementia with up to 10 years follow up: the gothenburg MCI study. *J. Alzheimer's Dis.* 44, 205–214. doi: 10.3233/JAD-141053

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6

Heister, D., Brewer, J. B., Magda, S., Blennow, K., and McEvoy, L. K. (2011). Predicting MCI outcome with clinically available MRI and CSF biomarkers. *Neurology* 77, 1619–1628. doi: 10.1212/WNL.0b013e3182343314

Holland, D., McEvoy, L. K., Desikan, R. S., and Dale, A. M. (2012). Enrichment and stratification for predementia Alzheimer disease clinical trials. *PLoS ONE* 7:e47739. doi: 10.1371/journal.pone.0047739

Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., and Martin, R. L. (1982). A new clinical scale for the staging of dementia. *Br. J. Psychiatry* 140, 566–572. doi: 10.1192/bjp.140.6.566

Klein, J. P., Houwelingen, H. C., Ibrahim, J. G., and Scheike, T. H. (eds.). (2013). *Handbook of Survival Analysis*. London: Chapman and Hall/CRC.

Lambert, J.-C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., et al. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* 45, 1452–1458. doi: 10.1038/ng.2802

Li, J.-Q., Tan, L., Wang, H.-F., Tan, M.-S., Tan, L., Xu, W., et al. (2016). Risk factors for predicting progression from mild cognitive impairment to Alzheimer's disease: a systematic review and meta-analysis of cohort studies. *J. Neurol. Neurosurg. Psychiatry* 87, 476–484. doi: 10.1136/jnnp-2014-310095

Marioni, R. E., Campbell, A., Hagenaars, S. P., Nagy, R., Amador, C., Hayward, C., et al. (2017). Genetic stratification to identify risk groups for Alzheimer's disease. *J. Alzheimer's Dis.* 57, 275–283. doi: 10.3233/JAD-161070

McEvoy, L. K., Fennema-Notestine, C., Roddey, J. C., Hagler, D. J., Holland, D., Karow, D. S., et al. (2009). Alzheimer disease: quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment. *Radiology* 251, 195–205. doi: 10.1148/radiol.2511080924

McEvoy, L. K., Holland, D., Hagler, D. J., Fennema-Notestine, C., Brewer, J. B., and Dale, A. M. (2011). Mild cognitive impairment: baseline and longitudinal structural MR imaging measures improve predictive prognosis. *Radiology* 259, 834–843. doi: 10.1148/radiol.11101975

Querfurth, H. W., and LaFerla, F. M. (2010). Alzheimer's disease. *N. Engl. J. Med.* 362, 329–343. doi: 10.1016/B978-0-323-28047-1.00047-0

Ridge, P. G., Mukherjee, S., Crane, P. K., and Kauwe, J. S. K. (2013). Alzheimer's disease: analyzing the missing heritability. *PLoS ONE* 8:e79771. doi: 10.1371/journal.pone.0079771

Ritchie, C. W., Molinuevo, J. L., Truyen, L., Satlin, A., Van der Geyten, S., and Lovestone, S. (2016). Development of interventions for the secondary prevention of Alzheimer's dementia: the European prevention of Alzheimer's Dementia (EPAD) project. *Lancet Psychiatry* 3, 179–186. doi: 10.1016/S2215-0366(15)00454-X

Roberts, R., and Knopman, D. S. (2013). Classification and Epidemiology of MCI. *Clin. Geriatr. Med.* 29, 753–772. doi: 10.1016/j.cger.2013.07.003

Sabuncu, M. R., Buckner, R. L., Smoller, J. W., Lee, P. H., Fischl, B., and Sperling, R. A. (2012). The association between a polygenic Alzheimer score and cortical thickness in clinically normal subjects. *Cereb. Cortex* 22, 2653–2661. doi: 10.1093/cercor/bhr348

Tan, C. H., Fan, C. C., Mormino, E. C., Sugrue, L. P., Broce, I. J., Hess, C. P., et al. (2018). Polygenic hazard score: an enrichment marker for Alzheimer's associated amyloid and tau deposition. *Acta Neuropathol.* 135, 85–93. doi: 10.1007/s00401-017-1789-4

Tan, C. H., Hyman, B. T., Tan, J. J. X., Hess, C. P., Dillon, W. P., Schellenberg, G. D., et al. (2017). Polygenic hazard scores in preclinical Alzheimer's disease. *Ann. Neurol.* 82, 484–488. doi: 10.1002/ana.25029

Vemuri, P., Wiste, H. J., Weigand, S. D., Shaw, L. M., Trojanowski, J. Q., Weiner, M. W., et al. (2009). MRI and CSF Biomarkers in normal, MCL and AD subjects: predicting future clinical change. *Neurology* 73, 287–293. doi: 10.1212/WNL.0b013e3181af79e5

Westman, E., Muehlboeck, J. S., and Simmons, A. (2012). Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion. *Neuroimage* 62, 229–238. doi: 10.1016/j.neuroimage.2012.04.056

Yu, J.-T., Tan, L., and Hardy, J. (2014). Apolipoprotein E in Alzheimer's disease: an update. *Annu. Rev. Neurosci.* 37, 79–100. doi: 10.1146/annurev-neuro-071013-014300

Conflict of Interest Statement: LM holds equity in CorTechs Labs, Inc. AD is a founder of and holds equity in CorTechs Labs, Inc., and serves on its Scientific Advisory Board. He is also a member of the Scientific Advisory Board of Human Longevity, Inc. (HLI), and receives research funding from General Electric Healthcare (GEHC). The terms of these arrangements have been reviewed and approved by the University of California, San Diego in accordance with its conflict of interest policies. OA has a patent SYSTEMS AND METHODS FOR IDENTIFYING POLYMORPHISMS pending.

The other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 18 June 2018
doi: 10.3389/fnins.2018.00397

Plin4-Dependent Lipid Droplets Hamper Neuronal Mitophagy in the MPTP/p-Induced Mouse Model of Parkinson's Disease

Xiaojuan Han[1,2†], Jialei Zhu[3], Xinlei Zhang[1], Qiqi Song[1], Jianhua Ding[3], Ming Lu[3], Sifan Sun[1,4] and Gang Hu[1,3*]*

[1] Department of Pharmacology, School of Medicine and Life Sciences, Nanjing University of Chinese Medicine, Nanjing, China, [2] Department of Traditional Chinese Medicine, The Affiliated Drum Tower Hospital of Nanjing University Medical School, Nanjing, China, [3] Jiangsu Key Laboratory of Neurodegeneration, Department of Pharmacology, Nanjing Medical University, Nanjing, China, [4] Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, China

OPEN ACCESS

Edited by:
Leo P. Sugrue,
University of California,
San Francisco, United States

Reviewed by:
Patrícia Maciel,
Escola de Medicina da Universidade do Minho, Portugal
Federico Herrera,
Instituto de Tecnologia Química e Biológica (ITQB-NOVA), Portugal

***Correspondence:**
Sifan Sun
sifansun@163.com
Gang Hu
ghu@njmu.edu.cn;
neuropha@njmu.edu.cn

[†]Co-first author

Specialty section:
This article was submitted to Neurodegeneration, a section of the journal Frontiers in Neuroscience

Received: 12 February 2018
Accepted: 23 May 2018
Published: 18 June 2018

Citation:
Han X, Zhu J, Zhang X, Song Q, Ding J, Lu M, Sun S and Hu G (2018) Plin4-Dependent Lipid Droplets Hamper Neuronal Mitophagy in the MPTP/p-Induced Mouse Model of Parkinson's Disease. Front. Neurosci. 12:397. doi: 10.3389/fnins.2018.00397

Epidemiological studies have shown that both lipid metabolism disorder and mitochondrial dysfunction are correlated with the pathogenesis of neurodegenerative diseases (NDDs), including Parkinson's disease (PD). Emerging evidence suggests that deposition of intracellular lipid droplets (LDs) participates in lipotoxicity and precedes neurodegeneration. Perilipin family members were recognized to facilitate LD movement and cellular signaling interactions. However, the direct interaction between Perilipin-regulated LD deposition and mitochondrial dysfunction in dopaminergic (DA) neurons remains obscure. Here, we demonstrate a novel type of lipid dysregulation involved in PD progression as evidenced by upregulated expression of Plin4 (a coating protein and regulator of LDs), and increased intracellular LD deposition that correlated with the loss of TH-ir (Tyrosine hydroxylase-immunoreactive) neurons in the MPTP/p-induced PD model mouse mesencephalon. Further, *in vitro* experiments showed that inhibition of LD storage by downregulating Plin4 promoted survival of SH-SY5Y cells. Mechanistically, reduced LD storage restored autophagy, leading to alleviation of mitochondrial damage, which in turn promoted cell survival. Moreover, the parkin-poly-Ub-p62 pathway was involved in this Plin4/LD-induced inhibition of mitophagy. These findings were further confirmed in primary cultures of DA-nergic neurons, in which autophagy inhibitor treatment significantly countermanded the ameliorations conferred by Plin4 silencing. Collectively, these experiments demonstrate that a dysfunctional Plin4/LD/mitophagy axis is involved in PD pathology and suggest Plin4-LDs as a potential biomarker as well as therapeutic strategy for PD.

Keywords: Plin4, lipid droplets, Parkinson's disease, DA neurons, mitophagy

Abbreviations: 3-MA, 3-Methyladenine; CNS, central nervous system; ER, endoplasmic reticulum; FFAs, free fatty acids; IHC, immunohistochemistry; LDH, lactate dehydrogenase; LDs, lipid droplets; MPTP, methyl-4-phenyl-1,2,3,6-tetrahydropyridine; NC, negative control; NDDs, neurodegenerative diseases; PD, Parkinson's disease; poly-Ub, poly-ubiquitin; SNpc, substantia nigra pars compacta; TEM, transmission electron microscopy; TH, tyrosine hydroxylase.

INTRODUCTION

Parkinson's disease (PD) is a progressive neurodegenerative disorder characterized by the preferential loss of dopaminergic neurons in the SNpc, affecting 2–3% of the population over the age of 65 (Poewe et al., 2017). The exact etiology and natural course of PD have yet to be fully clarified but involve dysfunction of numerous system-level processes, including mitochondrial function, calcium and dopamine homeostasis, neuroinflammation, and autophagy (Athauda and Foltynie, 2015; Ascherio and Schwarzschild, 2016; Poewe et al., 2017), highlighting the predominant role of mitochondrial dysfunction (Ryan et al., 2015; Onyango et al., 2017). Genome-wide association studies (GWAS) have identified many of the PD-associated genes such as PINK, PARKIN and DJ-1, which have been shown to either directly or indirectly play roles in mitochondrial homeostasis or mitophagy (Dawson and Dawson, 2017; Onyango et al., 2017). Moreover, recent studies have indicated that abolished mitophagy-mediated clearance of mtDNA and mtROS rendered impaired mitochondria as activators of the NLRP3 inflammasome to trigger neuroinflammation and promote DA neuronal loss (Yan et al., 2015; Zhong et al., 2016). Thus, excessive mitochondrial stress in response to genetic or environmental toxins may result in the death of neurons, and the capability to remove damaged mitochondria through mitophagy must be well controlled. Thus, the role that regulation of mitophagy plays in the pathogenesis of PD needs further exploration.

Recent reports revealed that LD-related lipotoxicity might participate in PD pathology (Wang et al., 2002; Outeiro and Lindquist, 2003; Schaffer, 2016). LDs are highly dynamic organelles that emerge from the ER membrane and serve as the intracellular sites for neutral lipid storage (Hashemi and Goodman, 2015). Accumulating evidence suggests that LDs play a much broader role in biology than previously indicated, including sequestering transcription factors, generating ligands for nuclear receptors, and regulating immunity (Welte, 2015). As to NDDs, α-synuclein, a pathogenic protein in PD, was reported to bind to LDs *in vitro* (Thiam et al., 2013), and subsequent *in vivo* studies confirmed that aggregated LDs in glia accelerated neurodegeneration in Drosophila (Bailey et al., 2015; Liu et al., 2015). Combined, these studies suggest that an LD abnormality may contribute to NDDs. Mechanistically, emerging evidence reveals an unexpected intimate association between LDs and other intracellular organelles, especially the ER, autophagic lysosomes and mitochondria, to affect their functions (Jaishy and Abel, 2016). Notably, LDs were reported to affect mitochondrial fusion dynamics, ensuring maximum oxidative metabolism and homeostasis (Rambold et al., 2015). Thus, the precise modulation of LDs is maintained to guarantee mitochondrial quality control.

Perilipin family members (Plin1-5), the surrounding proteins of LDs, are regarded as the most important regulator of LDs, facilitating LD movement and cellular signaling interactions (Kimmel and Sztalryd, 2016). Among them, Plin2 has been widely studied and proposed as a target for counteracting both metabolic and age-related diseases (Conte et al., 2016). Plin4 (S3-12), a perilipin normally expressed in heart and skeletal muscle, was reported to be involved in cardiac lipid accumulation (Chen et al., 2013). These studies collectively support the idea that perilipins may play a common role in pathological degenerative conditions. Although a few reports have indicated a relationship between LDs and α-synuclein or mitochondria, the detailed involvement and underlying mechanism of perilipin-regulated LDs in the pathogenesis of PD are not yet clarified.

In this study, using both cell and animal models and a combination of RNA-seq and LD-specific indicators, we illustrate that the Plin4/LD/mitophagy axis has a crucial role in neurodegeneration resulting from MPTP/MPP^{++} insult and indicate Plin4-LDs as a potential biomarker as well as therapeutic strategy for PD.

MATERIALS AND METHODS

Mice

Male C57BL/6 mice (4 months old) weighing 24–30 g were purchased from the Model Animal Research Center of Nanjing Medical University (Nanjing, China). All mice were harbored in the specific pathogen-free facility in Nanjing Medical University. The animals were maintained with free access to pellet food and water in plastic cages at 21 ± 2°C and kept on a 12 h light–dark cycle. To evaluate the lipid metabolic reaction and related genetic changes in the MPTP/p PD models, mice were randomly divided into two groups (n = 18/group) for subsequent model induction. The chronic MPTP intoxication protocol was similar to that described previously (Zhou et al., 2016), control mice were treated with saline. All animals were sacrificed 1 week after the final injection. At the endpoint, mice were anesthetized by sodium pentobarbital (50 mg/kg, i.p.), then perfused and sectioned, followed by western blotting (n = 6/group), IHC staining (n = 6/group), TEM (n = 3/group) and RNA-seq (n = 3/group) analyses as shown in the schedule (**Figure 1A**) with details showed in respective method sections. For the culture of primary midbrain neurons, pregnant mice with embryonic (E15-17) fetuses were used. The study was approved by the Animal Ethical and Welfare Committee of Nanjing Medical University. All animal welfare and experimental procedures were performed in accordance with the Guide for the Care and Use of Laboratory Animals (National Institutes of Health, United States) and the related ethical regulations of Nanjing Medical University. And all efforts were made to reduce the number of animals used and to minimize animal suffering.

Cell Culture

Authenticated SH-SY5Y cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM, 8115192, 32016001) with 10% FBS (Gibco, Cat 10100147) and 1% penicillin-streptomycin. Cells were cultured to a confluency of 70–80% in six-well or 24-well dishes and transfected with 1 μg of *Plin4* siRNA or NC siRNA (100 nM, Santa Cruz, CA, United States) in OptiMEM using Lipofectamine^{TM} 3000 Transfection Reagent (Invitrogen, Cat L3000-015) for 6 h. For the induction of neuronal damage, cells were incubated with ultrapure MPP^{+} (Sigma, 200 μM) for 24 h. For pharmacological measurements, the autophagy

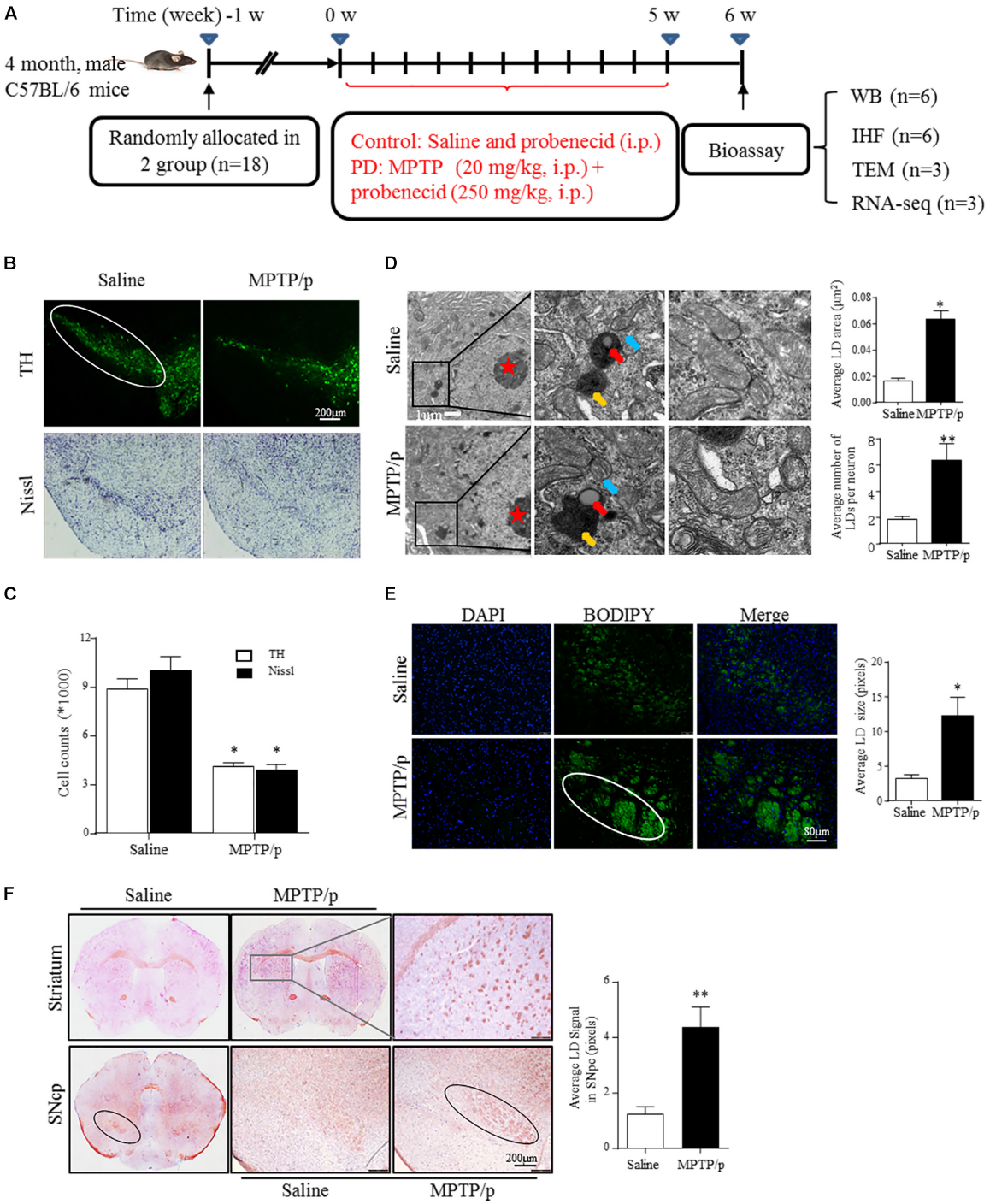

FIGURE 1 | Lipid droplets (LDs) accumulate in the mesencephalon of MPTP/p-treated mice. **(A)**: Diagram of the experimental design. Four-month-old male C57BL/6 mice were grouped and treated as indicated. Mice were transcardially perfused at the endpoint for bioassay. **(B,C)** A: TH and Nissl staining of the mouse brain sections. In a saline-injected control, there is a dense TH^+ and $Nissl^+$ network of cell bodies in the SNpc. After MPTP injection, there is a dramatic reduction in TH immunoreactivity and Nissl staining. Scale bar: 200 μm. B: Stereology of the loss of SNpc neurons in NC and MPTP-dosed mice. $^*P < 0.05$ vs. saline group, determined by one-way ANOVA ($n = 6$). **(D)** TEM of mesencephalon tissues comparing LD accumulation in neurons (with red stars indicating the nucleus) and distribution relative to organelles, as indicated. Middle frames with red arrows show lipid containing vesicles, yellow shows the autolysosome and blue shows mitochondria. Right frames displayed normal (up) and abnormal (down) mitochondria. **(E,F)** Tissues from **(B)** stained with BODIPY **(E)** and Oil Red O **(F)**. BODIPY showed more LD accumulation in SNpc (outlined in white) of PD model mice, in contrast to controls. **(F)** Whole-mount brain sections show LD accumulation in the SNpc (outlined in black), and abnormal intensely stained cross-sections of TH^+ axon morphology can be seen in the striatum (frame shows higher magnification) of PD model mice, in contrast to the controls. Scale bar as indicated. Data shown in **(D–F)**: quantitation of LD number and size. $^*P < 0.05$, $^{**}P < 0.01$ vs. saline group as determined by Student's *t*-test ($n = 4$).

inhibitor 3-MA (Sigma, Cat M-9281, 5 mM) was added to the cell culture medium 1 h before siRNA transfection. The cell or mitochondrial extracts were analyzed by Western blotting.

Primary mesencephalic neurons were obtained and cultured according to our previously described protocol (Xie et al., 2010). After siRNA transfection and incubation with 50 μM MPP^+ for 48 h, cells were rinsed carefully with PBS and fixed, followed by immunocytochemistry. As to quantification, the number of TH-ir neurons was counted in 10 randomly selected fields on a Nikon Optical TE2000-S inverted microscope. The values were normalized to that obtained from control cultures. The average number of TH-ir cells in control cultures ranged from 20 to 30 per field, with TH-ir cells making up approximately 5% of all cells in the primary culture. Each TH-ir cell process was traced from the soma to the end of the process and quantified by the measurement function of Image-Pro Plus 6.0.

Induction of the PD Model and Unbiased Stereology

The protocols for generating the MPTP/p-induced chronic PD mouse model and for unbiased stereology have been reported previously (Lu et al., 2014). Briefly, 20 mg/kg MPTP (Sigma, Cat M-0896) dissolved in saline was injected subcutaneously followed by 250 mg/kg DMSO-dissolved probenecid, which blocks the rapid clearance of the neurotoxin MPTP, injection intraperitoneally at 1 h interval every 3.5 days over a period of 5 weeks. Control mice were treated with saline and probenecid. At the endpoint, all animals were anesthetized and perfused, and brains were sectioned for Western blotting and IHC staining. For *in vivo* cell quantification studies, the number of TH^+ neurons and $Nissl^+$ neurons in the SNpc of the midbrain was assessed using the optical fractionator (Stereo Investigator 7, MBF Bioscience, Williston, VT, United States) as previously reported (Lu et al., 2014). All stereological analyses were performed under the 200× objective of an Olympus BX52 microscope (Olympus America Inc., Melville, NY, United States). The stereology was blinded to all genotype and treatment groups for each experiment.

Tissue Staining, Imaging and Quantification

For frozen samples, mice were perfused transcardially with 4% paraformaldehyde. Brains were extracted, post-fixed, dehydrated, embedded in OCT (Tissue-Tek), and cryosectioned at 30 μm per slice. For immunofluorescence, slides were incubated with the indicated primary antibodies at 4°C overnight, then washed and incubated in secondary fluorescent antibodies, followed by mounting in Prolong Gold Antifade with DAPI (Life Technologies, Cat P36931) before imaging. Images were observed and photographs were captured under a confocal microscope (Axiovert LSM510, Carl Zeiss Co., Germany). The integrated optical densities (IODs) were calculated using ImageJ by sampling of a 30 × 30 pixel area, and 36 images were captured from six consecutive mesencephalon sections. The values were reported as the average intensity above the background ± SD.

For Oil Red O staining, a working Oil Red O solution was generated by diluting a 3.5 mg/ml stock (in 100% isopropanol) (Sigma, Cat 0625) 3:2 with distilled water. This solution was incubated at room temperature for 30 min and filtered with Whatman paper before use. Sections were incubated in 60% isopropanol for 2 min, dried, and incubated in Oil Red O staining solution for 1 h at room temperature. Slides were rinsed in distilled water and counterstained with hematoxylin prior to mounting on Prolong glass slides. For LD staining in CNS sections, slides were submerged in PBS for 10 min and then incubated for 10 min in $BODIPY^{493/503}$ (Life Technologies, Cat D3922). The slides were then washed in PBS and immediately covered with Vectashield mounting medium with DAPI for later imaging on the same day. The LD staining signals were quantitatively analyzed using ImageJ as described above for the immunostaining signals.

Cell Staining With $BODIPY^{493/503}$ and MitoTracker Deep Red

Live cells were washed twice in PBS and incubated with 2 μg/ml $BODIPY^{493/503}$ (Life Technologies, Cat D3922) in PBS for 15 min at 37°C. For MitoTracker Deep Red staining, live cells were incubated with 0.5 μg/ml MitoTracker Deep Red (Invitrogen, Cat M22426) in PBS for 30 min at 37°C. After staining, the cells were washed in PBS and fixed in 3.5% PFA for 10 min. Then, the cells were washed and counterstained with Hoechst 33342 (Sigma, Cat B2261) for 10 min before being covered on glass slides for imaging. Images were observed and photographs were captured under an optical inverted fluorescence microscope (Nikon, TE2000-S).

Flow Cytometry Analysis

Mitochondrial membrane potential was measured by fluorescence levels upon staining with JC-1 (Invitrogen, Cat M7514) and MitoTracker Deep Red (Invitrogen, Cat M22426) at 0.5 μg/ml for 30 min at 37°C according to the manufacturer's instructions. Apoptosis of cells was assessed by staining cells with Annexin V/PI (Invitrogen, Cat V13242) at 37°C for 30 min according to the manufacturer's instructions. The cells were then washed with PBS and resuspended in cold PBS containing 1% FBS for flow cytometric analyses with Guava easyCyte System 8 (Millipore 25801, Hayward, CA, United States).

RNA-seq Analysis

Transcriptional profiling via RNA-seq analysis was conducted using the Illumina Hiseq kit according to the manufacturer's instructions. HTSeq v0.6.1 was used for quantification of gene expression level. Differential expression analysis of the two groups was performed using the DESeq R package (1.10.1). In detail, DESeq provide statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting *P*-values were adjusted using the Benjamini and Hochberg's approach for controlling the False Discovery Rate (FDR). Genes with an adjusted *P*-value < 0.05 found by DESeq were assigned as differentially expressed. As a result, a particular subset of 74

genes was generated as differentially expressed, with the summary statistical data shown in Supplementary Materials.

RNA Reverse-Transcription and Quantitative RT-PCR Analysis

Total RNA was extracted with Trizol reagent (Invitrogen Life Technologies, Carlsbad, CA, United States). Total RNA (1 μg) of each sample was reverse-transcribed into cDNA and amplified using a PrimeScript™ RT Master Mix (Takara, RR036A, Takara Biotechnology, China) according to the manufacturer's directions. RT-PCR was measured using a QuantiTect SYBR® Green PCR kit (Qiagen, Germany) with an ABI 7300 StepOne™ Fast Real-Time PCR System (Applied Biosystems, Foster City, CA, United States). The primer sequences used in this study are listed in the Supplementary Data. After the addition of primers and template DNA to the master, PCR thermal cycle parameters were as follows: 95°C for 3 min, 40 cycles of 60°C for 30 s and 95°C for 15 s, and a melting curve from 60 to 95°C to ensure amplification of a single product. In each sample the *GAPDH* gene was used as an endogenous control to normalize for differences in the amount of total RNA.

Western Blot, Transmission Electron Microscopy Analysis and Hoechst Staining

The protocols for the Western blot and TEM assays have been reported previously (Lu et al., 2014; Jia et al., 2017). Midbrain and cell protein lysates were quantified by Bradford assays (Bio-Rad, Hercules, CA, United States). Proteins were electrophoresed through a 10–15% SDS–polyacrylamide gel and transferred to PVDF membrane (Millipore, IPVH00010). The membranes were probed with the indicated primary antibodies followed by HRP-conjugated secondary antibodies. Signals were detected by enhanced chemiluminescence (ECL) Western blot detection reagents (Pierce, Rockford, IL, United States). The membranes were scanned and analyzed using an Image Quant LAS 4000 Imaging System (GE Healthcare, Stockholm, Sweden). The average blot intensities were calculated using ImageJ, and the values are reported as the average intensity above the background with β-actin used for normalization.

Antibody Details

Primary Antibodies

Plin4 (Novus, Cat 13776,1:500 for IF &1:1000 for WB); TH (Abcam, Cat ab6211,1:800 for IF); GFAP (Millipore, Cat MAB360,1:500 for IF); poly-Ub (CST, Cat 3936,1:500 for IF & 1:1000 for WB); Tom20 (CST, Cat 13929s,1:500 for IF & 1:1000 for WB);LC3B (CST, Cat 2775,1:500 for IF & 1:1000 for WB); p62 (CST, Cat 5114,1:500 for IF & 1:1000 for WB); Parkin(Abcam, Cat ab15954,1:800 for WB), Cytochrome C (CST, Cat 4272s,1:800 for WB), AIF(E-1) (Santa Cruz, Cat sc-13116,1:800 for WB), caspase-3 (CST, Cat 9662s,1:500 for WB), Beclin-1(CST, Cat 3738,1:1000 for WB), ATG7 (ABGENT, Cat AP1813a,1:800 for WB), β-actin (Santa Cruz, Cat SC-47778,1:1000 for WB). Secondary antibodies: Alexa F488 Donkey anti-mouse (Invitrogen, Cat A21202, 1:1000), Alexa Fluor 488 Goat anti-Rabbit (Invitrogen, Cat A11008,1:1000), Alexa Fluor 594 Donkey anti-Goat (Invitrogen, Cat A11058, A11008, A 11059,1:1000), HRP-conjugated secondary antibodies (Thermo, Cat 0031430, 31460, 31402, 1:1000).

Statistical Analysis

All data are presented as the means ± SEM and were collected and analyzed in a blinded manner. Statistical analysis was performed using Student's *t*-test or one-way analysis of variance (ANOVA) followed by the Holm–Sidak test (SigmaPlot 11.0). Two-way ANOVA was used when the genotype and treatment were considered as two independent variables. The tests used are indicated in the figure legends. In all studies, n indicates the number of samples per group, and cases in which P-values < 0.05 were considered statistically significant.

RESULTS

LDs Accumulate in the Mesencephalon of MPTP/p-Treated Mice

Emerging studies have documented unexpected LD accumulation accompanied by mitochondrial dysfunction that promotes neurodegeneration in Drosophila (Liu et al., 2015). To determine whether LD formation occurs in mammalian models and participates in the pathogenesis of PD, we generated the MPTP neurotoxin-induced PD mouse model for chronic mitochondrial dysfunction-associated loss of DA neurons using a previously described protocol (Lu et al., 2014). As shown by IHC staining, PD model mice were well established, with both TH^+ and $Nissl^+$ neurons in the SNpc showing a dramatic reduction in number (41.5 and 52.5%, respectively) in MPTP-treated mice compared with those in the saline-injected control group (**Figures 1B,C**).

Transmission electron microscopy indicated that MPTP injection induced a significant accumulation of LDs in the cytosol of neurons (indicated by red stars; morphologically seen with low and homogeneous electron density, little heterochromatin, and a large, round nucleolus for neuronal cell identification). LDs are easily identifiable in **Figure 1D**, similar to results reported previously (Singh et al., 2009), showing round, light-density structures not limited by a lipid bilayer membrane (Middle frame, red arrow), with homogenous amorphous content that was increased in both size and amount in response to MPTP/p stimuli. In contrast, LDs were rarely observed or were found to have a much smaller size in controls. Moreover, MPTP/p-induced LD accumulation was accompanied by mitochondrial damage (Right frame) and nearby autolysosomes (middle frame, yellow arrow). These findings suggest abnormal LD deposition in SNpc neurons that may correlate with mitochondrial damage in this PD model. For further confirmation, we performed LD-specific staining in brain sections and observed LD deposition (enhanced neutral lipid) in both the SNpc and striatum by using $BODIPY^{493/503}$ (**Figure 1E**) and Oil Red O (**Figure 1F**). LD-specific staining was especially dense in both the SNpc and striatum of injured animals, indicating much more LD

accumulation in regions where the bodies and axons of TH-ir neurons are located, whereas control animals showed diffuse staining (**Figures 1E,F**). These results suggested MPTP/p toxicity could induce mitochondrial damage with concomitant LD accumulation in the midbrain neurons of PD model mice.

Increased Plin4 Expression Is Positively Correlated With LD Accumulation in the Midbrain of MPTP/p-Treated Mice

Lipid droplets functions are regulated predominantly by the surrounding perilipin proteins of LDs (Kimmel and Sztalryd, 2016). To confirm whether and which perilipin may contribute to LD accumulation in the PD model mouse brain, we conducted the RNA-seq and subsequent quantitative PCR analysis of mesencephalon samples from saline- and MPTP/p-challenged mice. As quantitative PCR revealed, among all the increased genes (P-value < 0.05, details in Supplemental Materials), *Plin4* showed the highest overexpression in MPTP/p-treated mouse brain (**Figure 2A**). In contrast, the mRNA level of other perilipins was not elevated (**Figure 2B**). We also found a much higher *Plin4* level in midbrain and striatum, compared to that in cortex and hippocampus (Supplementary Figure S1), further supporting the relationship between *Plin4* and PD. Correspondingly, Western blotting also revealed a significant upregulation of Plin4 protein in the midbrain of PD model mice (**Figure 2C**). As shown in **Figure 2D**, co-immunostaining of Plin4 and different cell markers demonstrated that Plin4 mainly localized in TH^+ neurons but not in $GFAP^+$ glia. Indeed, MPTP/p challenge-induced elevation of Plin4 expression was further confirmed by IHF (**Figure 2E**), supporting the evidence that the LD-related Plin4 change indeed occurred in DA neurons. To confirm this result *in vitro*, SH-SY5Y cells, a representative cell line of dopaminergic neurons, were cultured and then stimulated with MPP^+ to mimic MPTP/p-induced neurological damage *in vivo*. Consistent with the *in vivo* findings, MPP^+ treatment promoted Plin4 expression and intracellular neutral lipid deposition as determined by $BODIPY^{493/503}$ staining (**Figure 2F**) in SH-SY5 cells. Together, these findings indicate increased Plin4 expression, rather than that of other perilipins, may account for LD accumulation and participate in PD pathogenesis.

Knockdown of *Plin4* Decreases LD Accumulation and Alleviates SH-SY5Y Cell Damage

To directly examine the role of Plin4 in LD accumulation due to MPP^+ stimuli *in vitro*, we transfected multiple siRNAs targeting *Plin4* or a scrambled control (SCR) into SH-SY5Y cell lines and observed that both *Plin4* mRNA (Supplementary Figure S2) and protein expression were decreased (**Figure 3A**). LDH release assay and Hoechst 33342 staining both indicated that the *Plin4* knockdown alone had no effect on cell viability (**Figures 3D,E**). By staining with the lipid-specific indicator Oil Red O and $BODIPY^{493/503}$, we further revealed that *Plin4* silencing indeed reversed MPP^+-induced intracellular LD accumulation (**Figures 3B,C**). Furthermore, to assess the functional capacity of LDs, SH-SY5Y cells were transfected with *siPlin4* or SCR plus MPP^+ treatment, and we found that Plin4 depleted cells were more tolerant to MPP^+ induced toxicity, evidenced by both Hoechst 33342 staining and LDH release assay (**Figures 3D,E**), consistent with reduced lipid storage within LDs. In addition to the Annexin V/PI staining indicated, the percentage of apoptotic cells (**Figure 3F**) further confirmed that cell damage was alleviated by *Plin4* silencing. Collectively, these results showed that *Plin4* inhibition decreased MPP^+-induced LD deposition and enhanced cell viability. Thus, *Plin4* is essential for LD deposition and accounts for cell damage upon toxic mitochondrial insult.

Knockdown of *Plin4* Reverses Mitochondrial Damage in the MPP^+ Cell Model

It has been reported that altered LDs could trigger dysfunction of many intracellular organelles, especially the mitochondria (Welte, 2015), and that this is mainly mediated by the surface coated protein, perilipin. Thus, we proposed that Plin4-mediated LD deposition could magnify mitochondrial stress, contributing to subsequent death-signals being released and damage initiated. To confirm this, we first examined whether Plin4 affects the subcellular distribution of LDs and/or leads to mitochondrial translocation by MPP^+ stimuli. As showed in **Figure 4A**, LDs displayed rare and diffuse cytoplasmic distribution in PBS-treated controls. In contrast, MPP^+ incubation caused LD deposition and localization with or adjacent to, damaged mitochondria. Moreover, *Plin4* deficiency could prevent this MPP^+-induced LD shifting. Then, to assess the functional association between *Plin4* and mitochondrial damage, the mitochondrial membrane potential was measured by JC-1 staining in SH-SY5Y cells. As **Figure 4B** indicated, the decrease in membrane potential induced by MPP^+ was alleviated by *Plin4* silencing. We also examined the expression of apoptosis-inducing factor (AIF) and Cytochrome C (cyto C), both of which translocate from the mitochondria into the cytoplasm in the early phases of mitochondrial damage to initiate the apoptotic proteolytic cascade by activating caspase-3. Western blotting showed more AIF and cyto C released into the cytoplasm upon MPP^+ stimulation (**Figures 4C,D**), which was partly reversed by *Plin4* deletion. For direct mitochondrial integrity assessment, we further observed the ultramicromorphological changes in SH-SY5Y cells. MPP^+ triggered accumulation of numerous swollen mitochondria in the SCR control cells, containing many highly damaged, electron-dense disrupted cristae mitochondria, which was attenuated by *Plin4* knockdown (**Figures 4E,F**). These results indicate that genetic silencing of *Plin4* alleviated MPP^+-induced mitochondrial damage in a dopaminergic cell line.

Downregulation of *Plin4* Activates Mitophagy in MPP^+-Treated SH-SY5Y Cells

Mitophagy, a quality control process, mediates the clearance of damaged ubiquitinated mitochondria, and mutations in genes related to mitophagic function, especially PINK1 and DJ-1, account for many autosomal recessive forms and

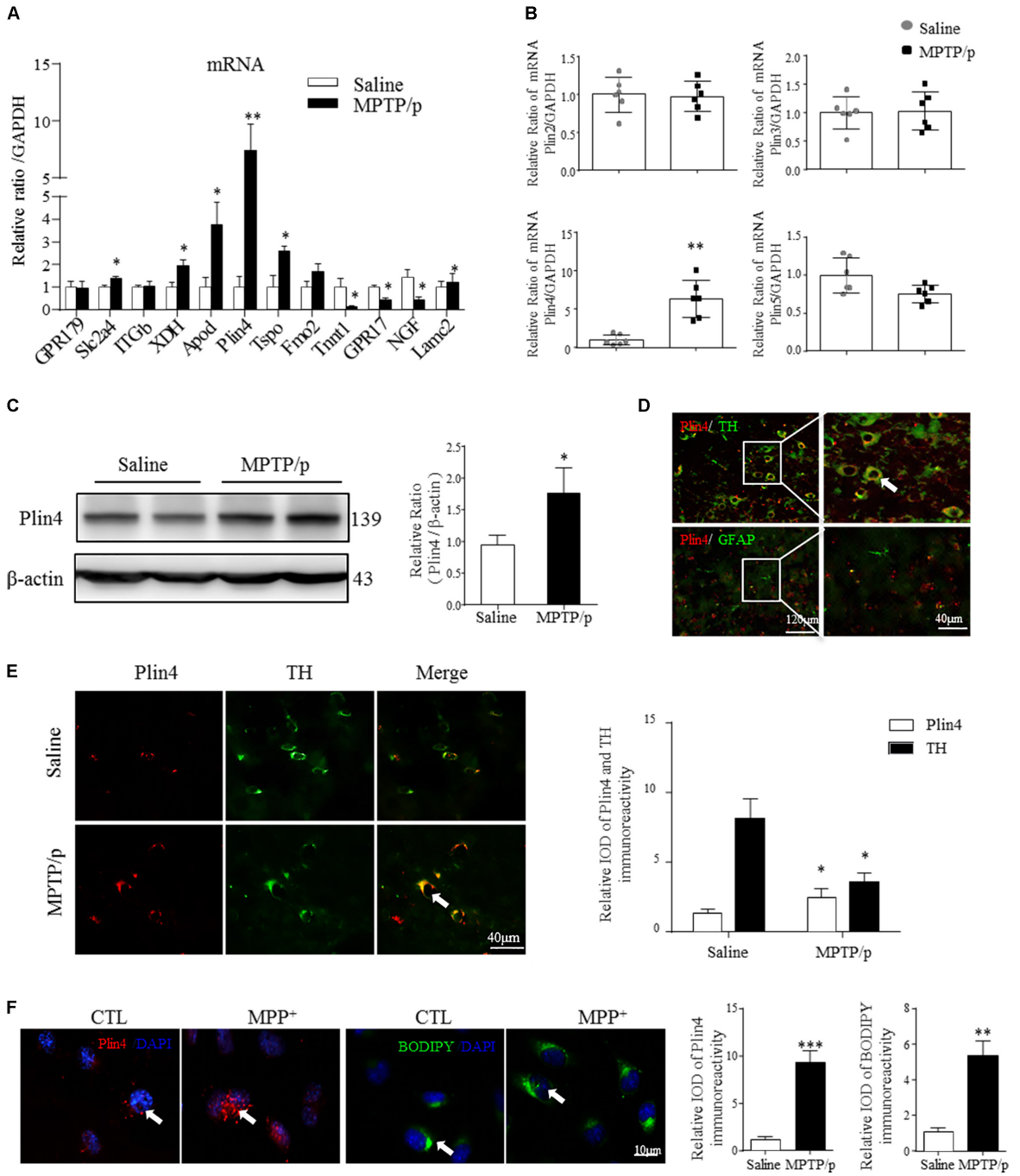

FIGURE 2 | Increased Plin4 is positively correlated with LD accumulation in the midbrain of MPTP/p-treated mice. **(A,B)** A: Quantitative RT-PCR analysis of RNA-seq indicated increased genes. B: RT-PCR analysis of Plin4 and other perilipins in mesencephalon samples from matched PD and NC mice. **(C)** Western blot analysis was performed to assess Plin4 expression *in vivo*. **(D,E)** For cellular location, the expression of Plin4 and TH/GFAP were analyzed by immunofluorescence of frozen brain sections. Arrows indicate neurons with coexpression of Plin4 and TH; the integrated optical density (IOD) of Plin4 and TH staining is presented on the right. **(F)** SH-SY5Y cells were stimulated with MPP^+ for 24 h followed by IF of Plin4 and $BODIPY^{493/503}$ staining (arrows marked), quantitation showed in right. Scale bar as indicated. Data are presented as the means ± SEM. Data in **(A,E)**: $^*P < 0.05$, $^{**}P < 0.01$ vs. the control group as determined by one-way ANOVA ($n = 4$). Data in **(B,C,F)**: $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ vs. the control group as determined by Student's *t*-test ($n = 6$ in **B,C**).

some sporadic cases of PD (Lazarou et al., 2015). As shown in **Figure 5A**, MPP^+ induced an increase in LC3-labeled vacuoles (namely, foci of LC3-II, the autophagosomal marker microtubule-associated protein 1 light chain 3) with a localization adjacent to mitochondria (as indicated by Tom20, a mitochondrial membrane protein) in SH-SY5Y cells. To

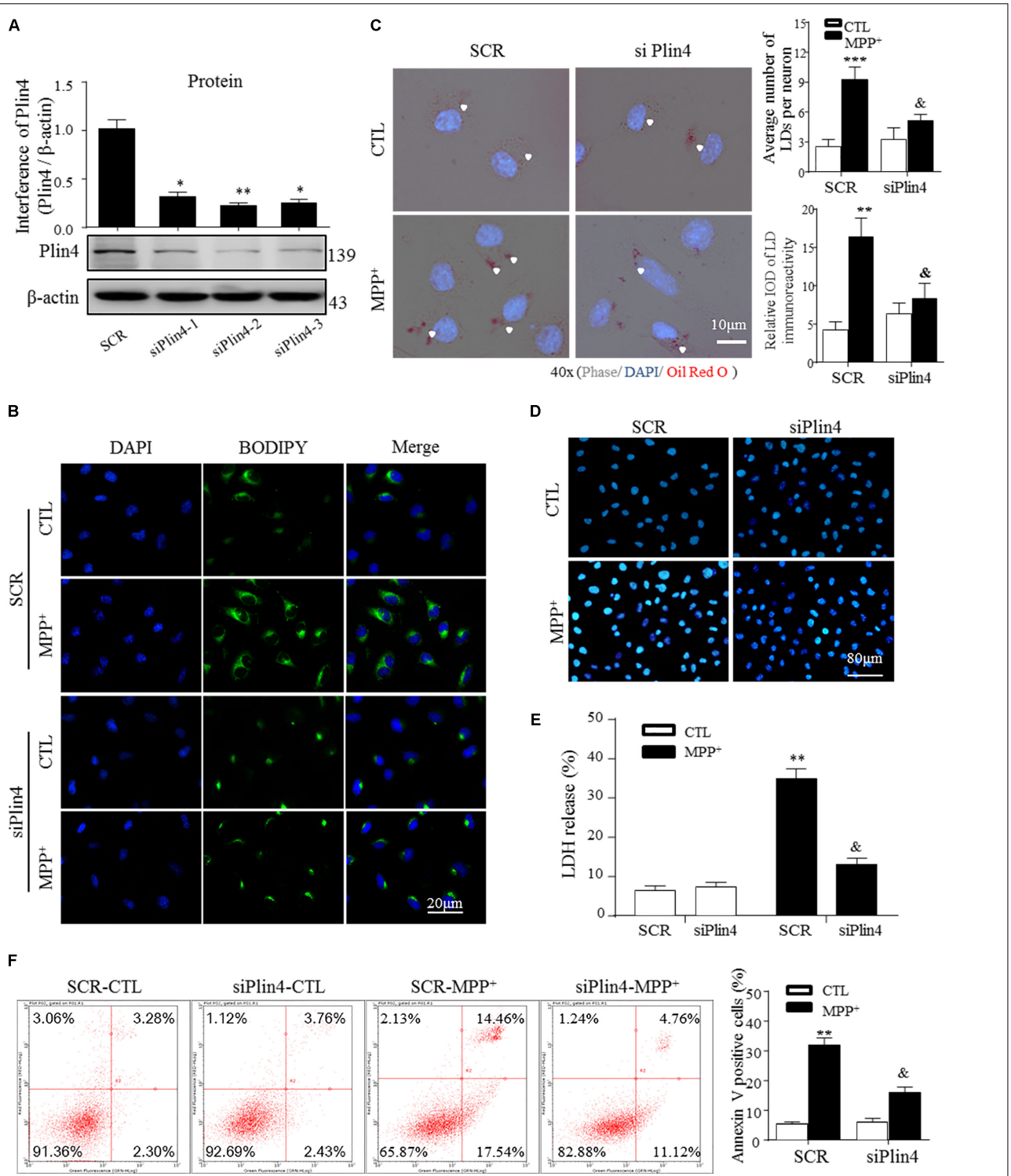

FIGURE 3 | Knockdown of *Plin4* decreases LD accumulation and alleviates SH-SY5Y cell damage. **(A)** SH-SY5Y cells were transduced with siRNA against *Plin4* or an SCR control. Western blot analysis was performed at 48 h post-transduction to assess for Plin4 suppression. $^{**}P < 0.01$, $^{***}P < 0.001$ vs. SRC as determined by one-way ANOVA followed by the Holm–Sidak test. **(B)** Cells described in **(A)** followed by CTL (PBS alone) or 200 μM MPP^+ stimulation for another 24 h. Live cells were stained with $BODIPY^{493/503}$ (2 μg/ml) then fixed and counterstaining with DAPI for fluorescence imaging. **(C)** Oil Red O staining of the cells described in **(B)**. Right: quantitation of LD number and integrated optical density (IOD). **(D)** *Plin4* knockdown and MPP^+ stimulation-induced cytotoxicity in SH-SY5Y cells were determined by Hoechst 33342 staining. **(E)** The cells described in **(B)** were subjected to the LDH release assay after 24 h of treatment. **(F)** Annexin V/PI staining indicated apoptosis of the cells described in **(E)**. Data in **(C,E,F)** are shown as mean ± SEM. $^{*}P < 0.05$, $^{**}P < 0.01$ vs. SCR-CTL, $^{\&}P < 0.05$ vs. SCR-MPP^+ as determined by two-way ANOVA.

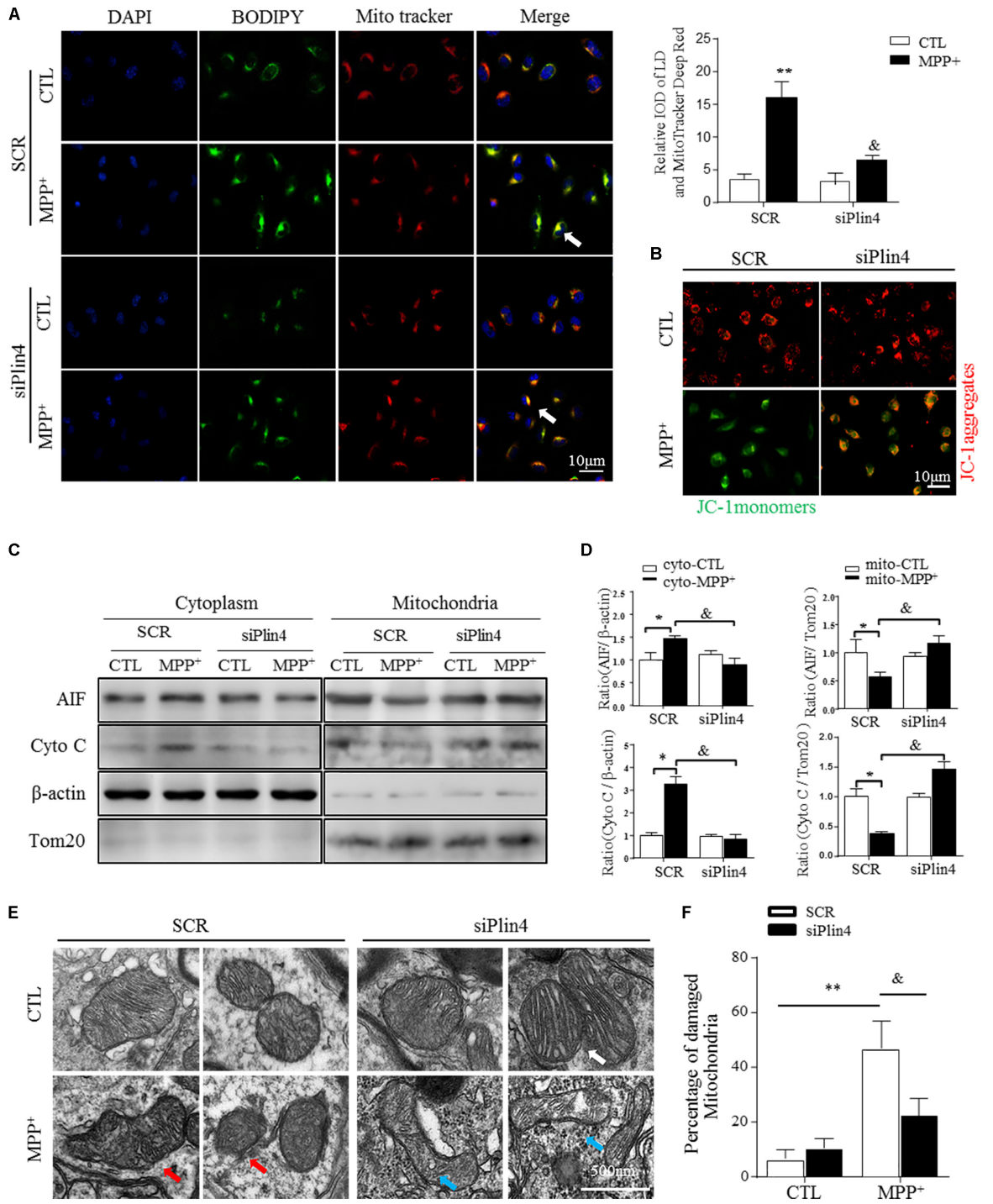

FIGURE 4 | Knockdown of *Plin4* reverses mitochondrial damage in MPP$^+$ cell model. **(A)** Intracellular distribution of LDs (BODIPY$^{493/503}$) and mitochondria (MitoTracker Deep Red) in MPP$^+$-stimulated SH-SY5Y cells were examined by fluorescence microscopy. Arrows indicate cells with colocalization of LDs and mitochondria; the integrated optical density (IOD) of BODIPY$^{493/503}$ and MitoTracker Deep Red immunofluorescence is presented on the right. **(B)** MPP$^+$-induced mitochondrial membrane potential changes in WT (SCR) or *Plin4*-deficient (siPlin4) SH-SY5Y cells were measured by JC-1 staining. Scale bar: 10 μm. **(C)** SH-SY5Y cells described in **(B)** were harvested, and then, mitochondria and cytosol were separated using a commercial kit and assessed for AIF and Cyto C expression. **(D)** Quantification of protein expression in **(C)**. β-actin and Tom20 were utilized as endogenous control genes for cytosol and mitochondria, respectively, and relative expression levels were determined by normalizing to those in the SCR-CTL samples. **(E)** Electron micrographs of mitochondria in WT (SCR) or *Plin4*-deficient (siPlin4) SH-SY5Y cells incubated with MPP$^+$ as described above. Shown are representative examples of normal (white arrow), partially damaged (blue arrow), and heavily damaged mitochondria (red arrow). Scale bar: 500 nm. **(F)** Quantification of damaged mitochondria in **(E)**. Data in **(A,D,F)** are shown as mean ± SEM. $^*P < 0.05$, $^{**}P < 0.01$ vs. SCR-CTL, $^{\&}P < 0.05$ vs. SCR-MPP$^+$ as determined by two-way ANOVA.

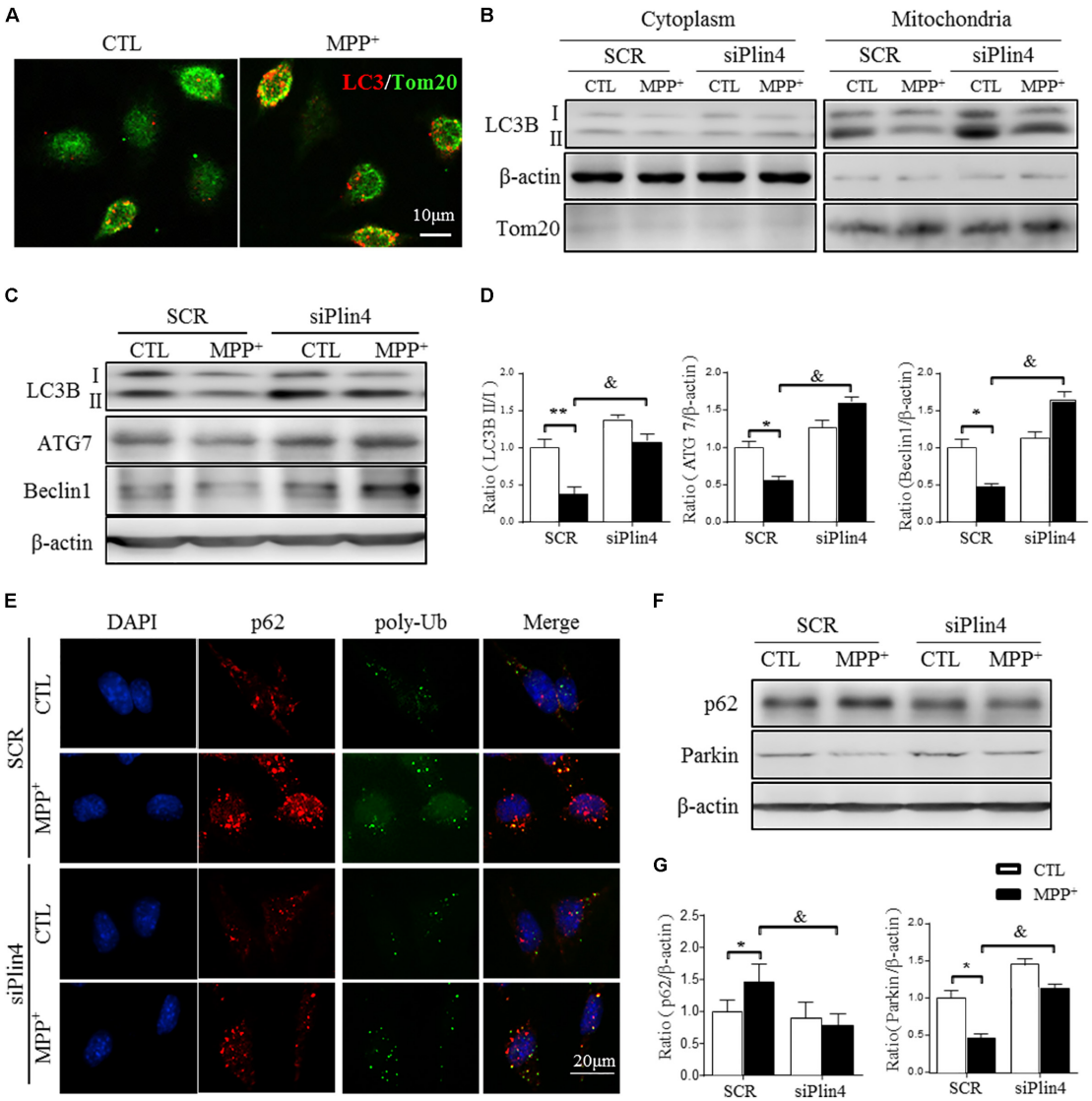

FIGURE 5 | Downregulation of *Plin4* activates autophagy in MPP$^+$-treated SH-SY5Y cells. **(A)** SH-SY5Y cells were treated with 200 μM MPP$^+$ for 24 h. LC3-II dot localization was detected by ICF with Tom20 as a mitochondrial marker. Scale bar: 10 μm. **(B)** SH-SY5Y cells were transfected with SCR or siPlin4 for 24 h, followed by MPP$^+$ incubation (200 μM, 24 h). Mitochondria and cytosol were then separated and assessed for LC3B expression. **(C,D)** C: Cells treated as in **(B)** were harvested, and proteins were collected for detection of LC3B, ATG7, and Beclin1 expression, with quantification shown in **(D)**. **(E)** Cells described in **(B)** were fixed, followed by ICF staining of p62 and poly-Ub sequestration and intracellular localization. Scale bar: 20 μm. **(F,G)** Cells described in **(C)** were harvested for detection of p62 and Parkin expression and relative quantification. Data in **(D,G)** are shown as the mean ± SEM. $^*P < 0.05$, $^{**}P < 0.01$ vs. SCR-CTL, $^{\&}P < 0.05$ vs. SCR-MPP$^+$ as determined by two-way ANOVA.

detect the participation of *Plin4* in mitophagy inhibition, we separated mitochondria from the cytoplasm and examined the expression of MAP1LC3B/LC3B-II, which is a cleaved MAP1LC3B-phosphatidylethanolamine conjugate and a general autophagosomal marker. We showed that Plin4 knockdown promoted mitophagy, as evidenced by more conversion of LC3-I to LC3-II in the mitochondria, which was previously inhibited by MPP$^+$ stimulation (**Figure 5B**). To further characterize the underlying autophagic mechanisms of *Plin4* in mitophagy inhibition, we further analyzed the expression of many autophagy-related proteins in mitochondria. As well as LC3B-II production, *Plin4* silencing also restored MPP$^+$-induced ATG7 and beclin1 inhibition (**Figures 5C,D**). These results indicated that MPP$^+$ blunted mitophagy could be rescued by *Plin4* silencing in SH-SY5Y cells.

During the autophagic process, p62 acts as a receptor protein that links LC3B with ubiquitinated substrates for clearance (Komatsu et al., 2012). Consistently, we found that MPP$^+$-induced p62 and poly-Ub sequestration was reduced by *Plin4* deficiency, as evidenced by both immunofluorescence (**Figure 5E**) and Western blotting (**Figures 5F,G**). Parkin, an identified mitophagy-related E3 ligase (Lazarou et al., 2015;

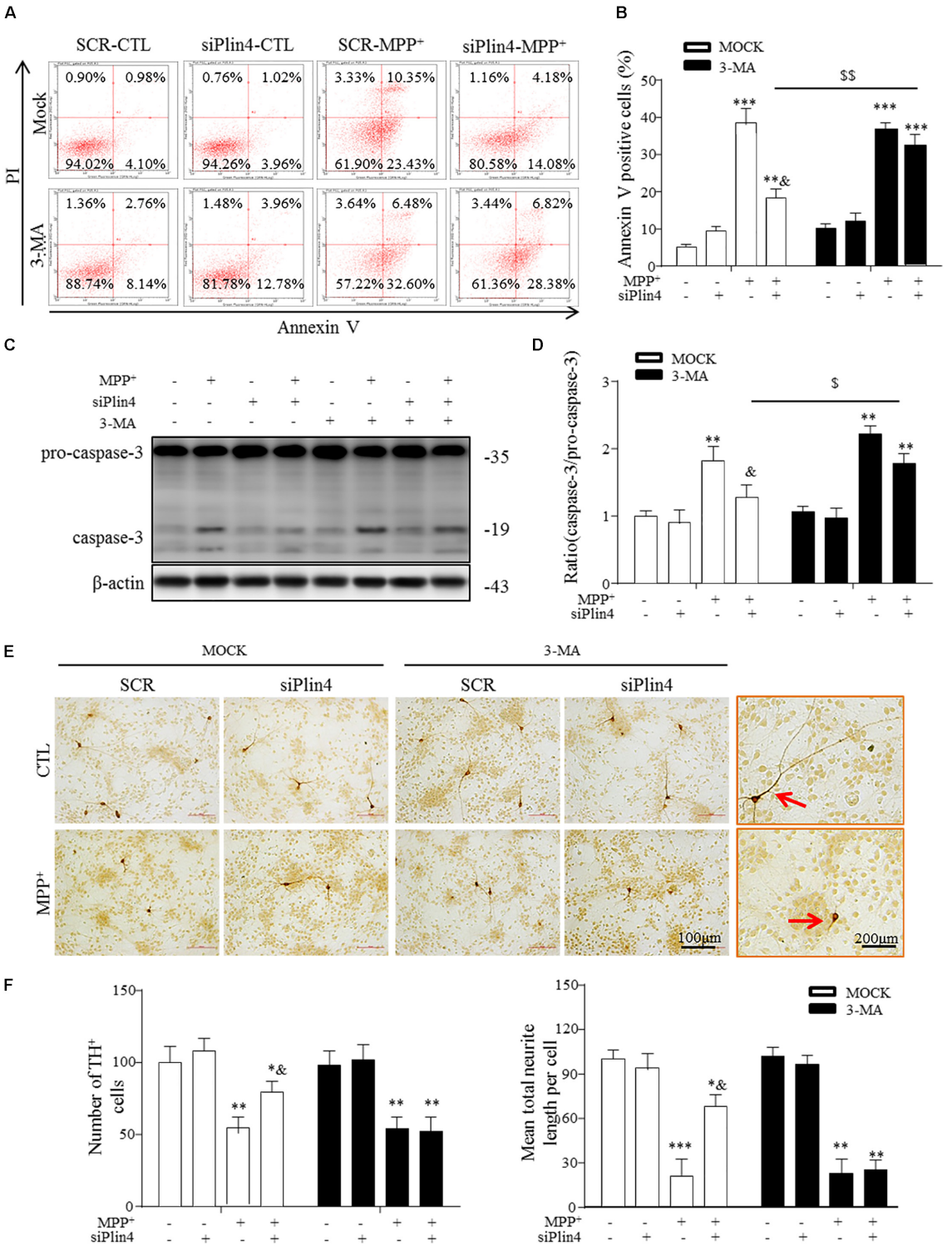

FIGURE 6 | 3-MA abolishes the protective effect of Plin4 deficiency on primary DA neurons. **(A,B)** A: SH-SY5Y cells were pretreated with PBS or 3-MA, followed by siRNA transfection 24 h before MPP^+ was given, and apoptosis was assessed by Annexin V/PI staining and flow cytometry. The data for flow cytometry analysis are presented as a percentage of the cell population by normalizing to MOCK-SCR-CTL samples. **(C,D)** Cells described in **(A)** were harvested and proteins were collected for the detection of caspase-3 activity by Western blot. **(E,F)** E: Primary cultured neurons of the mesencephalon suffered indicated treatments, followed by TH IHC detection; amplifications shown in the right frames for details. F: Representative number of TH^+ neurons and neurite length compared with SCR-NC. Scale bar as indicated. Data in **(B,D,F)** are shown as the mean ± SEM (n = 4–6). $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$ vs. SCR-NC-CTL; $^{\&}P < 0.05$ vs. SCR-NC-MPP^+; and $^{\$}P < 0.05$,$^{\$\$}P < 0.01$ vs. MOCK-NC-MPP^+ as analyzed by one-way ANOVA followed by the Holm–Sidak test.

Pickrell and Youle, 2015), is a key mediator of this p62-Ub mediated clearance of damaged mitochondrial and was also reversed by Plin4 deficiency (**Figures 5F,G**). Collectively, these results imply that inhibition of Plin4 promoted mitochondrial homeostasis by activating autophagic removal of damaged substrates in the PD cell model.

3-MA Abolishes the Protective Effect of *Plin4* Deficiency on Primary DA Neurons

Damaged mitochondria activate multiple signals, such as mitochondrial reactive oxygen species (mtROS), mtDNA and cardiolipin, promoting the mitochondrial apoptosis pathway, which represents the primary reason for the loss of TH-ir neurons in PD (Suliman and Piantadosi, 2016; Dawson and Dawson, 2017). Because our data indicate that Plin4 hindered autophagy, which accounted for MPP^+-induced mitochondrial damage, we speculated that promotion of the autophagic elimination of these danger signals by controlling *Plin4* might attenuate injuries triggered by MPTP toxicity. Therefore, we treated NC and Plin4-silenced SH-SY5Y cells with 3-MA, an autophagy inhibitor that can block the maturation and degradation of autophagy. Flow cytometry results using Annexin V/PI indicated that 3-MA pretreatment abolished the cell death-reducing effect of *Plin4* knockdown in MPP^+-stimulated SH-SY5Y cells (**Figures 6A,B**). These results were confirmed by immunoblot analysis of activation of caspase-3 (**Figures 6C,D**), an apoptosis-related protein that is proteolytically activated upon exposure to apoptotic stimuli. To provide more direct evidence linking DA neuronal deterioration, mitophagy and *Plin4*, primary mesencephalic neurons were cultured as described previously (Qiao et al., 2016), followed by MPP^+ stimulation to mimic the toxicity of MPTP/p *in vivo*. As shown in **Figures 6E,F**, MPP^+ decreased the number of TH-ir neurons by 43.4% and shortened neurite length by 65.7% compared with those in controls. In agreement with the results obtained in SH-SY5Y cells, MPP^+-induced damage in primary TH-ir neurons was also partially alleviated by siPlin4 transfection. Moreover, the neuroprotection contributed by Plin4 silencing was suppressed by 3-MA, supporting the idea that autophagy was involved in the *Plin4*-mediated deterioration of DA neurons.

DISCUSSION

The accumulation of lipid metabolites in nonadipose tissues causes lipotoxicity (Wymann and Schneiter, 2008) and is correlated with insulin resistance, type II diabetes mellitus, hepatic steatosis, and cardiovascular disease (Wymann and Schneiter, 2008; Gan et al., 2014; Schaffer, 2016). Lipotoxicity was found clinically relevant with NDDs, (Aviles-Olmos et al., 2013; Zou et al., 2015; Ascherio and Schwarzschild, 2016), although clear links have not been elucidated. Inside cells, lipids are stored within micelles known as LDs. Emerging evidence suggests the development of lipotoxicity is not solely due to the presence of FFAs in nonadipose tissues but is also due to the alteration of LD homeostasis (Krahmer et al., 2013; Thiam and Beller, 2017). Defects in LD biogenesis or turnover lead to perturbance of many metabolic pathways, further causing inflammation and mitochondrial and ER stress (Welte, 2015; Mukhopadhyay et al., 2017). Clinically, new evidence has emphasized the importance of LDs in cancers and aging-related diseases, as well as nervous system homeostasis (Bailey et al., 2015; Liu et al., 2015; Qiu et al., 2015; Conte et al., 2016). Recently, Liu et al. (2015) revealed that mutations in mitochondrial-related genes exhibit a common phenotype of LD accumulation, promoting the onset of neurodegeneration (ND) in Drosophila (Liu et al., 2015). However, a direct link between LD accumulation and NDD in mammals has not been established thus far.

In this study, via TEM and lipid-specific staining, we showed that MPTP/p-treated PD model mice exhibited LD accumulation in exactly the same brain region where DA neuronal loss occurs. Bailey et al. reported similar LD changes occurred in glial cells to preserve neuroblast proliferation under hypoxic conditions in Drosophila (Bailey et al., 2015). In our mouse model, both TEM and IHF co-immunostaining of TH/GFAP with Plin4 consistently revealed that LDs mainly clustered in neurons, with less accumulation in astrocytes. Whether LDs may perform some glial-specific function requires further exploration. Thus, our *in vivo* study suggests a strong correlation between LD deposition and mitochondrial stress in a specific MPTP/p-induced mouse model of PD.

Our results fit within an emerging theme in which primary injuries coordinately induce both lipid synthesis processes and a disturbance in mitochondrial homeostasis, which cooperate in reducing cell viability, representing a common change in many energy metabolism-related diseases, including NDDs (Qiu et al., 2015; Thiam and Beller, 2017). Specifically, evidence indicates LD accumulation resulting from mitochondrial dysfunction precedes physical and histological changes in NDD (Liu et al., 2015) and also protects neural stem cells from hypoxia damage during development in Drosophila (Bailey et al., 2015). Thus, in the case of energy barriers predominantly in CNS, LDs, the novel but incompletely defined organelles, may represent the primary and unique way for the organisms to solve the problem of balancing energy supply and cellular homeostasis. However, once the compensation cannot be satisfied or the stimuli sustained, LDs may subsequently change to be harmful and aggravate the injuries by inducing mitochondrial stress, ER stress or autolysosome dysfunction (Welte, 2015), as happened in our study.

Lipid droplets are highly dynamic organelles surrounded by a phospholipid monolayer and several proteins (Welte, 2015). The most abundant and well-characterized LD coating proteins belong to the so-called PAT family, comprising five members (1–5) collectively known as Perilipins (Kimmel and Sztalryd, 2016), which have physiological roles in facilitating storage of neutral lipids within LDs and regulating the intracellular interactions with other organelles. Here, with the combined use of RNA-seq and RT-PCR, we identified upregulated expression of Plin4 accompanied by LD accumulation in the MPTP/p-treated mouse midbrain. Plin4 is normally found decorating nascent LDs in the cytosol, and its inactivation has been reported to affect heart function (Chen et al., 2013). Moreover, genetic silencing of Plin4 significantly

ameliorated LD storage and promoted the survival of neurons *in vitro*. Hence, the molecular mechanisms underlying Plin4-controlled LD deposition are likely to be conserved in mice and may contribute to the pathogenesis of PD. The clarification of this question may help to provide novel explanations for the theory of lipotoxicity in PD.

Mitochondrial disability is a common feature of both familial and sporadic PD, as well as toxin-induced parkinsonism (Newmeyer and Ferguson-Miller, 2003; Ryan et al., 2015; Suliman and Piantadosi, 2016). Recent findings have further revealed mitophagy dysfunction to be clinically relevant, with abnormal lipid metabolism being suggested to relate to PD pathogenesis. Here, we showed that Plin4 silencing conferred a reduction in LD deposition and in turn resulted in prevention of MPP^+-induced mitochondrial damage as evidenced by mitochondrial membrane potential collapse, mitochondrial fragmentation, and mitophagy inhibition. Thus, our findings fit with the newly identified mitochondria-related roles of LDs (Welte, 2015).

In addition to their interactions with mitochondria, LDs also interact with autolysosomes, with several studies demonstrating a bidirectional relationship between the two (Jaishy and Abel, 2016). By examining the linkage between LD disturbance and autophagic signaling, we showed that Plin4 silencing-driven autophagy and p62-ubiqutin mediate the autophagic clearance of damaged mitochondria, which in turn ameliorated deterioration in both SH-SY5Y cells and primary cultured DA neurons. We present our conclusions based on the following observations: Plin4 deficiency conferred apoptosis inhibition and caspase-3 inactivation, which were significantly blocked by the autophagy inhibitor 3-MA.

CONCLUSION

In conclusion, by exploring the role of LDs *in vitro* and in a mouse PD model *in vivo*, we demonstrated that Plin4-dependent LD deposition in TH-ir neurons contributed to DA neuronal loss in MPTP/p-treated mice. Mechanistically, excessive accumulation of LDs may trigger mitochondria-impaired mitophagy, further resulting in subsequent neurodegenerative damage. Thus, a dysfunctional Plin4/LD/mitophagy axis is clarified to be involved in PD pathophysiology, indicating that Plin4-LD changes in the brain may be a promising biomarker as well as therapeutic target for PD.

The main limitation of this work is the use of a chemical insult-based model of PD, which cannot mimic the whole spectrum of PD pathogenesis. Replicating these findings in different PD models would help to verify the possibility of rendering Plin4-LDs as therapeutic targets for PD.

AUTHOR CONTRIBUTIONS

GH and ML p designed the research. XH performed co-transfection for the western blotting assays, neuropathological study, RT-PCR and drafted the manuscript. JZ performed the immunohistochemistry study and statistical analysis. JZ and SS contributed to the statistical analysis. XZ and QS helped to perform the western blotting assays and contributed materials/analysis tools. JD and SS discussed the project and gave valuable suggestions to this project. All authors read and approved the final manuscript.

FUNDING

This work was supported by grants from the National Science and Technology Major Project (No. 2017ZX09101003-003-007), the National Natural Science Foundation of China (Nos. 81630099, 81603091, and 81773706), and the key project of Natural Science Foundation of the Higher Education Institutions of Jiangsu Province (No. 15KJA310002).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2018.00397/full#supplementary-material

REFERENCES

Ascherio, A., and Schwarzschild, M. A. (2016). The epidemiology of Parkinson's disease: risk factors and prevention. *Lancet Neurol.* 15, 1257–1272. doi: 10.1016/S1474-4422(16)30230-7

Athauda, D., and Foltynie, T. (2015). The ongoing pursuit of neuroprotective therapies in Parkinson disease. *Nat. Rev. Neurol.* 11, 25–40. doi: 10.1038/nrneurol.2014.226

Aviles-Olmos, I., Limousin, P., Lees, A., and Foltynie, T. (2013). Parkinson's disease, insulin resistance and novel agents of neuroprotection. *Brain* 136, 374–384. doi: 10.1093/brain/aws009

Bailey, A. P., Koster, G., Guillermier, C., Hirst, E. M. A., MacRae, J. I., Lechene, C. P., et al. (2015). Antioxidant role for lipid droplets in a stem cell niche of *Drosophila*. *Cell* 163, 340–353. doi: 10.1016/j.cell.2015.09.020

Chen, W., Chang, B., Wu, X., Li, L., Sleeman, M., and Chan, L. (2013). Inactivation of Plin4 downregulates Plin5 and reduces cardiac lipid accumulation in mice. *Am. J. Physiol. Endocrinol. Metab.* 304, E770–E779. doi: 10.1152/ajpendo.00523.2012

Conte, M., Franceschi, C., Sandri, M., and Salvioli, S. (2016). Perilipin 2 and age-related metabolic diseases: a new perspective. *Trends Endocrinol. Metab.* 27, 893–903. doi: 10.1016/j.tem.2016.09.001

Dawson, T. M., and Dawson, V. L. (2017). Mitochondrial mechanisms of neuronal cell death: potential therapeutics. *Annu. Rev. Pharmacol. Toxicol.* 57, 437–454. doi: 10.1146/annurev-pharmtox-010716-105001

Gan, L. T., Van Rooyen, D. M., Koina, M., Mccuskey, R. S., and Teoh, N. C. (2014). Hepatocyte free cholesterol lipotoxicity results from JNK1-mediated mitochondrial injury and is HMGB1 and TLR4-dependent. *J. Hepatol.* 61, 1376–1384. doi: 10.1016/j.jhep.2014.07.024

Hashemi, H. F., and Goodman, J. M. (2015). The life cycle of lipid droplets. *Curr. Opin. Cell Biol.* 33, 119–124. doi: 10.1016/j.ceb.2015.02.002

Jaishy, B., and Abel, E. D. (2016). Lipids, lysosomes, and autophagy. *J. Lipid Res.* 57, 1619–1635. doi: 10.1194/jlr.R067520

Jia, M., Li, C., Zheng, Y., Ding, X., Chen, M., Ding, J., et al. (2017). Leonurine exerts antidepressant-like effects in the chronic mild stress-induced depression

model in mice by inhibiting neuroinflammation. *Int. J. Neuropsychopharmacol.* 20, 886–895. doi: 10.1093/ijnp/pyx062

Kimmel, A. R., and Sztalryd, C. (2016). The perilipins: major cytosolic lipid droplet-associated proteins and their roles in cellular lipid storage, mobilization, and systemic homeostasis. *Annu. Rev. Nutr.* 36, 471–509. doi: 10.1146/annurev-nutr-071813-105410

Komatsu, M., Kageyama, S., and Ichimura, Y. (2012). p62/SQSTM1/A170: physiology and pathology. *Pharmacol. Res.* 66, 457–462. doi: 10.1016/j.phrs.2012.07.004

Krahmer, N., Farese, R. J., and Walther, T. C. (2013). Balancing the fat: lipid droplets and human disease. *EMBO Mol. Med.* 5, 973–983. doi: 10.1002/emmm.201100671

Lazarou, M., Sliter, D. A., Kane, L. A., Sarraf, S. A., Wang, C., Burman, J. L., et al. (2015). The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. *Nature* 524, 309–314. doi: 10.1038/nature14893

Liu, L., Zhang, K., Sandoval, H., Yamamoto, S., Jaiswal, M., Sanz, E., et al. (2015). Glial lipid droplets and ROS induced by mitochondrial defects promote neurodegeneration. *Cell* 160, 177–190. doi: 10.1016/j.cell.2014.12.019

Lu, M., Sun, X., Qiao, C., Liu, Y., Ding, J., and Hu, G. (2014). Uncoupling protein 2 deficiency aggravates astrocytic endoplasmic reticulum stress and nod-like receptor protein 3 inflammasome activation. *Neurobiol. Aging* 35, 421–430. doi: 10.1016/j.neurobiolaging.2013.08.015

Mukhopadhyay, S., Schlaepfer, I. R., Bergman, B. C., Panda, P. K., Praharaj, P. P., Naik, P. P., et al. (2017). ATG14 facilitated lipophagy in cancer cells induce ER stress mediated mitoptosis through a ROS dependent pathway. *Free Radic. Biol. Med.* 104, 199–213. doi: 10.1016/j.freeradbiomed.2017.01.007

Newmeyer, D. D., and Ferguson-Miller, S. (2003). Mitochondria: releasing power for life and unleashing the machineries of death. *Cell* 112, 481–490. doi: 10.1016/S0092-8674(03)00116-8

Onyango, I. G., Khan, S. M., and Bennett, J. J. (2017). Mitochondria in the pathophysiology of Alzheimer's and Parkinson's diseases. *Front. Biosci.* 22, 854–872. doi: 10.2741/4521

Outeiro, T. F., and Lindquist, S., (2003). Yeast cells provide insight into alpha-synuclein biology and pathobiology. *Science* 5651, 1772–1775. doi: 10.1126/science.1090439

Pickrell, A. M., and Youle, R. J. (2015). The roles of PINK1, parkin, and mitochondrial fidelity in Parkinson's disease. *Neuron* 85, 257–273. doi: 10.1016/j.neuron.2014.12.007

Poewe, W., Seppi, K., Tanner, C. M., Halliday, G. M., Brundin, P., Volkmann, J., et al. (2017). Parkinson disease. *Nat. Rev. Dis. Primers* 3:17013. doi: 10.1038/nrdp.2017.13

Qiao, C., Yin, N., Gu, H. Y., Zhu, J. L., Ding, J. H., Lu, M., et al. (2016). Atp13a2 deficiency aggravates astrocyte-mediated neuroinflammation via NLRP3 inflammasome activation. *CNS Neurosci. Ther.* 22, 451–460. doi: 10.1111/cns.12514

Qiu, B., Ackerman, D., Sanchez, D. J., Li, B., Ochocki, J. D., Grazioli, A., et al. (2015). HIF2 -dependent lipid storage promotes endoplasmic reticulum homeostasis in clear-cell renal cell carcinoma. *Cancer Discov.* 5, 652–667. doi: 10.1158/2159-8290.CD-14-1507

Rambold, A. S., Cohen, S., and Lippincott-Schwartz, J. (2015). Fatty acid trafficking in starved cells: regulation by lipid droplet lipolysis, autophagy, and mitochondrial fusion dynamics. *Dev. Cell* 32, 678–692. doi: 10.1016/j.devcel.2015.01.029

Ryan, B. J., Hoek, S., Fon, E. A., and Wade-Martins, R. (2015). Mitochondrial dysfunction and mitophagy in Parkinson's: from familial to sporadic disease. *Trends Biochem. Sci.* 40, 200–210. doi: 10.1016/j.tibs.2015.02.003

Schaffer, J. E. (2016). Lipotoxicity: many roads to cell dysfunction and cell death: introduction to a thematic review series. *J. Lipid Res.* 57, 1327–1328. doi: 10.1194/jlr.E069880

Singh, R., Kaushik, S., Wang, Y., Xiang, Y., Novak, I., Komatsu, M., et al. (2009). Autophagy regulates lipid metabolism. *Nature* 458, 1131–1135. doi: 10.1038/nature07976

Suliman, H. B., and Piantadosi, C. A. (2016). Mitochondrial quality control as a therapeutic target. *Pharmacol. Rev.* 68, 20–48. doi: 10.1124/pr.115.011502

Thiam, A. R., Antonny, B., Wang, J., Delacotte, J., Wilfling, F., Walther, T. C., et al. (2013). COPI buds 60-nm lipid droplets from reconstituted water-phospholipid-triacylglyceride interfaces, suggesting a tension clamp function. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13244–13249. doi: 10.1073/pnas.1307685110

Thiam, A. R., and Beller, M. (2017). The why, when and how of lipid droplet diversity. *J. Cell Sci.* 130, 315–324. doi: 10.1242/jcs.192021

Wang, L., Schuster, G. U., Hultenby, K., Zhang, Q., Andersson, S., and Gustafsson, J. A. (2002). Liver X receptors in the central nervous system: from lipid homeostasis to neuronal degeneration. *Proc. Natl. Acad. Sci. U.S.A.* 99, 13878–13883. doi: 10.1073/pnas.172510899

Welte, M. A. (2015). Expanding roles for lipid droplets. *Curr. Biol.* 25, R470–R481. doi: 10.1016/j.cub.2015.04.004

Wymann, M. P., and Schneiter, R. (2008). Lipid signalling in disease. *Nat. Rev. Mol. Cell Biol.* 9, 162–176. doi: 10.1038/nrm2335

Xie, J., Duan, L., Qian, X., Huang, X., Ding, J., and Hu, G. (2010). KATP channel openers protect mesencephalic neurons against MPP+-induced cytotoxicity via inhibition of ROS production. *J. Neurosci. Res.* 88, 428–437. doi: 10.1002/jnr.22213

Yan, Y., Jiang, W., Liu, L., Wang, X., Ding, C., Tian, Z., et al. (2015). Dopamine controls systemic inflammation through inhibition of NLRP3 inflammasome. *Cell* 160, 62–73. doi: 10.1016/j.cell.2014.11.047

Zhong, Z., Umemura, A., Sanchez-Lopez, E., Liang, S., Shalapour, S., Wong, J., et al. (2016). NF-κB restricts inflammasome activation via elimination of damaged mitochondria. *Cell* 164, 896–910. doi: 10.1016/j.cell.2015.12.057

Zhou, Y., Lu, M., Du, R. H., Qiao, C., Jiang, C. Y., Zhang, K. Z., et al. (2016). MicroRNA-7 targets Nod-like receptor protein 3 inflammasome to modulate neuroinflammation in the pathogenesis of Parkinson's disease. *Mol. Neurodegener.* 11:28. doi: 10.1186/s13024-016-0094-3

Zou, Y. M., Liu, J., Tian, Z. Y., Lu, D., and Zhou, Y. Y. (2015). Systematic review of the prevalence and incidence of Parkinson's disease in the People's Republic of China. *Neuropsychiatr. Dis. Treat.* 11, 1467–1472. doi: 10.2147/NDT.S85380

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 16 July 2018
doi: 10.3389/fnins.2018.00476

Insulin-Like Growth Factor Binding Protein 2 Is Associated With Biomarkers of Alzheimer's Disease Pathology and Shows Differential Expression in Transgenic Mice

Luke W. Bonham[1,2†], Ethan G. Geier[1†], Natasha Z. R. Steele[1], Dominic Holland[3], Bruce L. Miller[1], Anders M. Dale[3,4,5], Rahul S. Desikan[2] and Jennifer S. Yokoyama[1] for the Alzheimer's Disease Neuroimaging Initiative[‡]*

[1] Department of Neurology, Memory and Aging Center, University of California, San Francisco, San Francisco, CA, United States, [2] Department of Radiology and Biomedical Imaging, University of California, San Francisco, San Francisco, CA, United States, [3] Department of Neurosciences, University of California, San Diego, San Diego, CA, United States, [4] Department of Psychiatry, University of California, San Diego, San Diego, CA, United States, [5] Department of Radiology, University of California, San Diego, San Diego, CA, United States

OPEN ACCESS

Edited by:
Kempuraj Duraisamy, University of Missouri, United States

Reviewed by:
Auriel Willette, National Institute on Aging (NIA), United States
Linda Zhang, Universidad Politécnica de Madrid (UPM), Spain

***Correspondence:**
Jennifer S. Yokoyama
jennifer.yokoyama@ucsf.edu

† These authors have contributed equally to this work.

Specialty section:
This article was submitted to Neurodegeneration, a section of the journal Frontiers in Neuroscience

Received: *27 February 2018*
Accepted: *25 June 2018*
Published: *16 July 2018*

Citation:
Bonham LW, Geier EG, Steele NZR, Holland D, Miller BL, Dale AM, Desikan RS and Yokoyama JS for the Alzheimer's Disease Neuroimaging Initiative (2018) Insulin-Like Growth Factor Binding Protein 2 Is Associated With Biomarkers of Alzheimer's Disease Pathology and Shows Differential Expression in Transgenic Mice. Front. Neurosci. 12:476. doi: 10.3389/fnins.2018.00476

There is increasing evidence that metabolic dysfunction plays an important role in Alzheimer's disease (AD). Brain insulin resistance and subsequent impairment of insulin and insulin-like growth factor (IGF) signaling are associated with the neurodegenerative and clinical features of AD. Nevertheless, how the brain insulin/IGF signaling system is altered in AD and the effects of these changes on AD pathobiology are not well understood. IGF binding protein 2 (IGFBP-2) is an abundant cerebral IGF signaling protein and there is early evidence suggesting it associates with AD biomarkers. We evaluated the relationship between protein levels of IGFBP-2 with cerebrospinal fluid (CSF) biomarkers and neuroimaging markers of AD progression in 300 individuals from across the AD spectrum. CSF IGFBP-2 levels were correlated with CSF tau levels and brain atrophy in non-hippocampal regions. To further explore the role of *IGFBP2* in tau pathobiology, we evaluated the expression of *IGFBP2* in different human and mouse brain cell types and brain tissue from two transgenic mouse models: the P301L-tau model of tauopathy and TASTPM model of AD. We observed significant differential expression of *IGFBP2* in both transgenic mouse models relative to wild-type mice in cortex but not in hippocampus. In both humans and mice, *IGFBP2* is most highly expressed in astrocytes. Taken together, our findings suggest that IGFBP-2 may be linked to tau pathology and provides further evidence for a relationship between metabolic dysregulation and neurodegeneration. Our results also raise the possibility that this relationship may extend beyond neurons.

Keywords: IGFBP-2, Alzheimer's disease, CSF, neuroimaging, tau

[‡]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

INTRODUCTION

Alzheimer's disease (AD) is the most common cause of dementia, affecting more than 37 million people worldwide (Matthews, 2010). The pathological hallmarks of AD are insoluble extracellular amyloid beta plaques and intracellular neurofibrillary tangles (NFTs) containing aggregates of hyperphosphorylated tau (p-tau) protein (Perl, 2010; Querfurth and LaFerla, 2010). This pathology is associated with neuronal cell loss and synaptic injury that leads to the characteristic memory loss, cognitive impairment, and behavioral changes observed in AD patients (Querfurth and LaFerla, 2010). An increasing number of studies suggest that alterations in brain metabolic processes play an important role in AD pathogenesis, with brain insulin resistance recognized as an important feature of AD in both patients and in post-mortem tissue (Salkovic-Petrisic and Hoyer, 2007; Querfurth and LaFerla, 2010; Takeda et al., 2010; Bomfim et al., 2012; Talbot et al., 2012; Stanley et al., 2016).

Insulin and insulin-like growth factor (IGF) signaling in the brain regulate neuronal growth, repair, and synaptic maintenance (Stockhorst et al., 2004; van Dam and Aleman, 2004), and play an important role in learning and memory (Zhao et al., 2004). Reduced sensitivity to insulin and IGF signals in AD are observed across multiple levels of cell signal response, including reduced insulin receptor (IR) sensitivity, inhibition of secondary messengers (i.e., IR substrate-1), and decreased IR and IGF receptor (IGFR) expression (Watson and Craft, 2003; Rivera et al., 2005; Steen et al., 2005; Holscher and Li, 2010; de la Monte, 2012; Talbot et al., 2012; Freiherr et al., 2013; Stanley et al., 2016). Markers of insulin resistance are elevated in hippocampus, a region of the brain that has high levels of IRs and is affected in AD (Frölich et al., 1998; Talbot et al., 2012; Bedse et al., 2015). Furthermore, insulin sensitivity restoration, insulin, and IGF therapy have been suggested to improve cognitive performance and memory function in healthy humans, individuals with mild cognitive impairment (MCI), and AD patients (Reger et al., 2008; Bomfim et al., 2012; Bedse et al., 2015), protect neurons from amyloid-induced toxicity in primary cell culture studies (Mattson, 1997), and prevent tau hyperphosphorylation in animal models (Deng et al., 2009; Chen et al., 2013). Thus, in the context of insulin and IGF resistance, neurons may be more vulnerable to damage and death resulting from the pathological changes underlying AD.

IGF-I and IGF-II are growth factors secreted by tissues throughout the body including the brain, and are effectively stored outside of cells in complex with IGF binding proteins (IGFBPs) (Holly and Perks, 2006). Interaction with IGFBPs slows IGF clearance and regulates the activity of IGF on cells through a variety of mechanisms (Russo et al., 2005). Importantly, several studies have observed altered levels of IGFs and IGFBPs in the plasma and cerebrospinal fluid (CSF) of AD patients, further suggesting that the neuroprotective and synaptic maintenance effects elicited by IGF signaling may be altered in AD (Tham et al., 1993; Vardy et al., 2007; Salehi et al., 2008; Hertze et al., 2014; Åberg et al., 2015). While the relationship between plasma levels of IGFs and IGFBPs and pathological features of AD have been explored to some extent (Toledo et al., 2013; Lane et al., 2017; Mclimans et al., 2017), knowledge of how plasma and CSF IGFs as well as IGFBPs contribute to AD pathogenesis remains in its early stages with previous reports focusing on cross-sectional analyses of neuroimaging data or longitudinal analyses of specific candidate regions. Further, studies of CSF IGFBPs in AD conflict, with some reporting significant differences in levels of IGFBPs (Salehi et al., 2008; Hertze et al., 2014; Lane et al., 2017; Mclimans et al., 2017) while others report no difference (Åberg et al., 2015).

In this study, we investigated the relationship between CSF IGFBP-2 and multiple *in vivo* markers of AD pathology to expand upon recent findings suggesting that IGFBP-2 plays a role in AD progression and pathogenesis (Lane et al., 2017; Mclimans et al., 2017). To better understand how IGFBP-2 may impact AD pathogenesis, we further utilized gene expression data from transgenic mouse models of tauopathy and AD along with cell type specific expression from human and mouse to assess the relevance of IGFBP-2 dysregulation to neurodegeneration.

MATERIALS AND METHODS

Participant Description

This study utilized samples from 300 individuals recruited for participation in the Alzheimer's Disease Neuroimaging Initiative (ADNI) study with CSF measurements of IGFBP-2 as well as amyloid, tau, and p-tau available. At baseline, 89 were cognitively normal older adults (CN), 145 individuals were diagnosed with MCI, and 66 were clinically diagnosed with AD. Two-hundred and seventy-six of these individuals had at least two T1-weighted MR images available. The cohort is well-characterized and has been used in previously published studies (Desikan et al., 2013, 2014; Bonham et al., 2016). Clinical severity of symptoms in the MCI and AD groupings was measured using the Clinical Dementia Rating Scale Sum of Boxes (CDR-SB) Score (Morris, 1993) and Mini Mental State Exam (MMSE) (Folstein et al., 1975). A clinician diagnosed each participant using a structured protocol that utilized clinical judgment and neuropsychological tests. Briefly, controls were required to have normal memory function on the Logical Memory II subscale of the Wechsler Memory Scale – Revised (Wechsler, 1987), an MMSE score greater than 24, CDR total score equal to 0, and clinical determination that the individual was not significantly impaired in cognitive function or activities of daily living. Individuals with MCI were required to have abnormal memory function on the Logical Memory II subscale of the Wechsler Memory Scale – Revised, an MMSE greater than 24, CDR total score equal to 0.5, and clinical determination that the individual's general cognition and functional performance was not impaired enough to make a diagnosis of AD. Finally, individuals with AD were required to have abnormal memory function on the Logical Memory II subscale of the Wechsler Memory Scale – Revised, an MMSE between 20 and 26, CDR total score equal to 0.5 or 1.0, and judgment by a clinician that the individual met NINCDS/ADRDA criteria for probable AD (McKhann et al., 1984). Informed and written consent was obtained from all study participants and the University of

California, San Francisco institutional review board approved all aspects of this study.

CSF Biomarker Measurements

The AlzBio3 Luminex xMAP immunoassay (Innogenetics, Ghent, Belgium) was used to measure CSF amyloid $\beta_{1\text{-}42}$ (amyloid), total tau (t-tau), p-tau_{181p} (p-tau) as described previously (Shaw et al., 2009; Kang et al., 2012). This method uses monoclonal antibodies specific for amyloid, t-tau, and p-tau. The monoclonal antibodies are chemically bonded to color-coded beads along with analyte-specific detector antibodies. Baseline CSF IGFBP-2 levels were measured using the Human DiscoveryMAP panel developed by Rules Based Medicine (Myriad RBM; Austin, TX, United States). The Human DiscoveryMAP panel is commercially available and measures a collection of metabolic, lipid, inflammatory, and other AD-relevant indicators. At the time this panel was used in the ADNI cohort, IGFBP-2 was the only IGF-related analyte in the panel. A full list of the measured metabolites is available through Myriad RBM. The CSF measurements in the immunoassay panel were processed and normalized according to previously described methods (Craig-Schapiro et al., 2011; Siuciak, 2011). Briefly, Myriad RBM used a Luminex 100 instrument for the measurements and analyzed the resulting data using proprietary software. The ADNI staff checked analyte distributions for normality using Box-Cox analyses and, if needed, log10 transformed the data to achieve an approximately normal distribution.

Genotyping and Gene Expression Data

APOE status in the ADNI cohort was determined using DNA extracted by Cogenics (now Beckman Coulter Inc., Pasadena, CA, United States) from a 3 mL aliquot of EDTA blood.

We evaluated *IGFBP2* expression using AD and tau transgenic mouse model data from mouseac (Matarin et al., 2015)[1]. Briefly, microarray gene expression data was collected from three brain regions (cortex, hippocampus, and cerebellum) from wild-type, TASTPM (TAS10 × TPM AD mouse models; APPswe × PS1.M1466V), and P301L-tau transgenic mice. Brain tissue samples were at 2, 4, 8, and 18 months of age and raw expression levels were normalized using a $\log_2$ transformation; all samples were quantile normalized together.

To better understand the cell type-specific expression of *IGFBP2*, we utilized two publicly available RNA sequencing expression datasets examining several cell-types commonly found in the central nervous system (CNS). For additional details on sample processing and cohort characteristics, please see Zhang et al. (2014, 2016) and Bennett et al. (2016).

Neuroimaging Data

One thousand one-hundred and sixteen T1-weighted MRI scans were processed using a quantitative volume and surface-based analysis technique which automatically segments scans into regions-of-interest (ROI) (Fischl et al., 2002; Desikan et al., 2006). The MRI scans were checked for quality and corrected for spatial distortion. All MRI scans were processed using Quarc (Quantitative Anatomical Regional Change), a modified version of the FreeSurfer pipeline designed to accurately estimate longitudinal changes in brain structure (Fennema-Notestine et al., 2007; Mcevoy et al., 2009; Holland and Dale, 2011; Holland et al., 2012). Cortical and subcortical ROIs were delineated using previously described automated parcellation and segmentation methods (Fischl et al., 2002; Desikan et al., 2006). The techniques used to estimate longitudinal sub-regional change for serial MRI scans are previously described (Holland and Dale, 2011). Briefly, Quarc utilizes non-linear registration of serial MR images to generate a deformation field that aligns both large and small structures with high fidelity. Volumetric changes are estimated as a percent change from the deformation field within a specified ROI. Quarc has been shown to be more a more sensitive measure of change over time compared to other measures of longitudinal brain atrophy such as the longitudinal FreeSurfer pipeline, TBM, and BSI (Holland et al., 2012). Quarc has been utilized extensively and has been shown to correlate closely with biomarkers of clinical progression (Desikan et al., 2011, 2013, 2014). We examined all 34 cortical regions of interest in the Desikan Killiany Atlas (Desikan et al., 2006) along with hippocampus and amygdala. For each region of interest, the change rate in the right and left structures was averaged.

Statistical Analysis

Demographic Comparisons

Discrete and continuous demographic variables were compared across diagnostic groups using chi-squared and ANOVA analyses, respectively.

Cross-Sectional CSF t-tau and p-tau Analyses

Linear models were used to test for an association between IGFBP-2 and t-tau and IGFBP-2 with p-tau. We controlled for age, sex, education, CDR-SB score, and *APOE* ε4 carrier status.

Neuroimaging Analyses

Linear mixed effects models were used to assess the relationship between IGFBP-2 levels and longitudinal gray matter atrophy controlling for baseline and time interactions of age, sex, education, baseline CDR-SB score, and *APOE* ε4 carrier status.

We used the following linear mixed effect model:

$$\Delta ROI\ Volume = \beta_0 + \beta_1 \Delta t + \beta_2 IGFBP\text{-}2 * \Delta t + \beta_3 \text{Age} * \Delta t + \beta_4\ \text{Sex} * \Delta t + \beta_5\ \text{Education} * \Delta t + \beta_6\ \text{CDR-SB} * \Delta t + \beta_7 APOE\varepsilon 4 * \Delta t + e$$

For neuroanatomical regions that showed volume change significantly predicted by IGFBP-2 levels only, we also assessed whether IGFBP-2 levels were associated with atrophy

[1] www.mouseac.org

TABLE 1 | Demographic information for participants included in the analysis.

	CN	MCI	AD	*p*-Value
N	89	145	66	NA
APOE ε4 Carrier (%)	23.5%	48.3%	56.1%	<0.001
Sex (% female)	50.6%	33.1%	28.8%	<0.01
Age (years)	75.7 ± 5.5	75.0 ± 7.2	74.9 ± 7.7	NS
Education (years)	15.6 ± 3.0	16.0 ± 3.0	15.0 ± 3.0	NS
CDR-SB	0.03 ± 0.1	1.56 ± 0.9	4.3 ± 1.6	<0.001
Aβ42 (pg/mL)	208.0 ± 52.9	160.7 ± 48.6	141.6 ± 35.6	<0.001
t-tau (pg/mL)	68.5 ± 26.8	104.2 ± 52.3	119.8 ± 54.6	<0.001
p-tau (pg/mL)	24.6 ± 12.9	36.0 ± 15.6	41.4 ± 20.5	<0.05
IGFBP-2 (ng/mL)	101.6 ± 18.0	104.8 ± 19.2	103.1 ± 18.8	NS

Summary statistics for participants. Demographic, genetic, and biomarker data is summarized by diagnostic category. APOE ε4 carrier includes those with 1 or 2 APOE ε4 alleles. CDR-SB, Clinical Dementia Rating Sum of Boxes. Two-tailed p-values were from ANOVA (continuous traits) or chi-square (categorical values) tests by sex, gene carrier status. CN, normal control; MCI, mild cognitive impairment; AD, Alzheimer's disease.

independent of baseline t-tau levels by adding the relevant terms to the original mixed effects model as follows:

$$\Delta ROI\ Volume = \beta_0 + \beta_1 \Delta t + \beta_2 IGFBP\text{-}2 * \Delta t + \beta_3\ \text{Age} * \Delta t + \beta_4\ \text{Sex} * \Delta t + \beta_5, \text{Education} * \Delta t + \beta_6 CDR\text{-}SB * \Delta t + \beta_7 APOE\ \varepsilon 4 * \Delta t + \beta_8 t\text{-}tau * \Delta t + e$$

The results of these analyses were used in statistical mediation analyses. We used the coefficients to perform the Aroian test as described by Preacher and Hayes (2004).

Gene Expression Analyses

ANOVA was used to determine whether *IGFBP2* expression varied between wild-type and tau transgenic mice in hippocampus and cortex. We chose not to analyze cerebellar expression because this region is generally spared in AD and the tau transgenic mouse models we used do not display cerebellar pathology.

RESULTS

Cohort Description

Data from 300 individuals identified as CN, MCI, or AD were included in this study (**Table 1**). The cohort was balanced with respect to age and education but differed by sex ($p = 7.35 \times 10^{-3}$). As expected, there were significant differences by diagnosis for *APOE* ε4 distribution, CSF amyloid, CSF t-tau, and CSF p-tau. CSF IGFBP-2 levels did not differ by diagnosis. For the 276 individuals with neuroimaging data, the demographic and biomarker composition was similar to the full cohort. The average number of scans per participant across the entire cohort was about 4 (CN: 3.97 ± 0.9, MCI: 4.40 ± 1.1, AD: 3.4 ± 0.8) with an average follow-up time per participant of 2.1 years (CN: 2.39 ± 0.8, MCI: 2.16 ± 0.7, AD: 1.73 ± 0.6). A histogram depicting the timing of follow-up of scans relative to the baseline visit is provided in Supplementary Figure 1.

CSF IGFBP-2 Is Associated With CSF t-tau and p-tau Levels

Across the entire cohort, IGFBP-2 was significantly associated with t-tau (β = 0.65 ± 0.15, $p = 2.41 \times 10^{-5}$) and p-tau (β = 0.17 ± 0.05, $p = 1.61 \times 10^{-3}$) in CSF, with higher levels of IGFBP-2 associated with higher levels of t-tau and p-tau (**Figure 1** and **Table 2**). Within subgroups, the association between CSF t-tau and IGFBP-2 was significant after correction for multiple testing in CN only (β = 0.58 ± 0.16, $p = 2.44 \times 10^{-4}$), with MCI (β = 0.54 ± 0.23, $p = 0.02$) and AD (β = 0.79 ± 0.40, $p = 0.05$) having *p*-values above $p = 0.017$ (Supplementary Table 1). For CSF p-tau, there were fewer observations available and IGFBP-2 was significant in MCI only (β = 0.18 ± 0.07, $p = 7.56 \times 10^{-3}$). However, the direction of the estimated effect in both CN (β = 0.08 ± 0.08, $p = 0.32$) and AD (β = 0.15 ± 0.15, $p = 0.31$) was consistent with the MCI grouping (Supplementary Table 1). There were no significant associations between IGFBP-2 levels and measures of CSF amyloid.

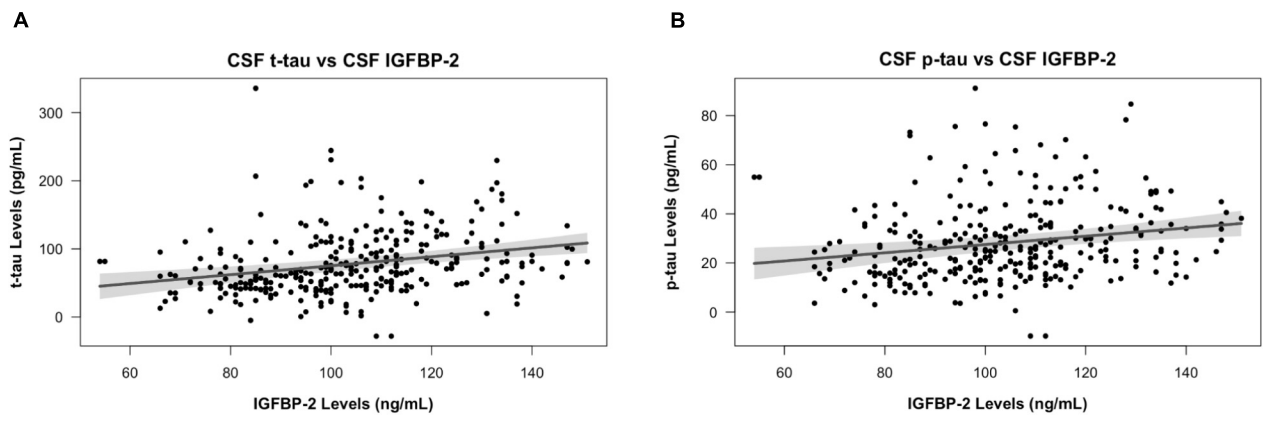

FIGURE 1 | IGFBP-2 is associated with t-tau and p-tau in CSF. CSF t-tau **(A)** and p-tau **(B)** levels are plotted against CSF IGFBP-2 levels. IGFBP-2 levels are quality controlled and transformed as previously described (Siuciak, 2011). Higher levels of IGFBP-2 are associated with higher levels of t-tau and p-tau. The plotted points are partial residuals with 95% confidence bands provided in shading.

TABLE 2 | Regression analysis results for predictors of cross sectional CSF p-tau and t-tau values in full cohort and by disease and meta-analysis groups.

Outcome	Variable	Estimate ± SE	*p*-Value
t-tau	Age	−0.63 ± 0.40	0.11
	Sex	18.33 ± 5.75	1.60×10^{-3}
	CDR-SB	6.26 ± 1.50	3.89×10^{-5}
	APOEε4 status	23.22 ± 5.40	2.31×10^{-5}
	Education	0.28 ± 0.88	0.75
	IGFBP-2	0.65 ± 0.15	2.41×10^{-5}
p-tau	Age	−0.28 ± 0.14	0.04
	Sex	2.40 ± 2.0	0.23
	CDR-SB	1.82 ± 0.52	5.30×10^{-4}
	APOEε4 status	9.40 ± 1.87	8.59×10^{-7}
	Education	0.06 ± 0.31	0.84
	IGFBP-2	0.17 ± 0.05	1.61×10^{-3}

IGFBP-2 is associated with t-tau and p-tau levels in CSF. Regression models used in cross-sectional CSF analyses of t-tau and p-tau are summarized. CSF IGFBP-2 was significantly associated with CSF t-tau and CSF p-tau. Higher levels of CSF IGFBP-2 were associated with higher levels of CSF t-tau and p-tau. The beta estimate (estimate) and accompanying standard error (SE) reflect the adjusted effect of each independent variable as a predictor of t-tau or p-tau. For all disease groups, the linear statistical model included as independent variables: age, sex, clinical disease rating sum of boxes (CDR-SB) APOE ε4 carrier status, education, and IGFBP-2. All tests were two-tailed.

CSF IGFBP-2 Is Associated With Brain Atrophy in AD-Associated Regions

We next tested whether participants' baseline CSF IGFBP-2 levels were associated with longitudinal volume change in all 34 Desikan Killiany cortical ROIs along with hippocampus and amygdala. At a raw $p < 0.05$, there were significant associations between CSF IGFBP-2 and atrophy in parahippocampal, entorhinal, inferior temporal, temporal pole, superior temporal, fusiform, isthmus cingulate, precuneus, rostral anterior cingulate, middle temporal, corpus callosum, caudal anterior cingulate, medial orbitofrontal, lateral occipital, and lateral orbitofrontal regions (**Figure 2A**). Additional details on the effect size and *p*-value for all regions are presented in Supplementary Table 2. After correction for multiple testing, CSF IGFBP-2 was significantly associated with atrophy in parahippocampal ($\beta = -0.30 \pm 0.06$, $p = 9.76 \times 10^{-5}$), entorhinal ($\beta = -0.31 \pm 0.06$, $p = 1.15 \times 10^{-3}$), inferior temporal ($\beta = -0.26 \pm 0.05$, $p = 3.85 \times 10^{-3}$), and temporal pole ($\beta = -0.32 \pm 0.12$, $p = 3.83 \times 10^{-3}$) regions (**Figure 2B**). All effects were consistent with greater CSF IGFBP-2 levels predicting greater atrophy over time.

Using mediation analysis, we found statistical evidence to suggest that CSF t-tau levels could partially explain the effects of IGFBP-2 on brain volume. We used the Aroian test to statistically assess whether the relationship between CSF IGFBP-2 and CSF t-tau mediated atrophy in parahippocampal, entorhinal, temporal pole, and inferior temporal regions. CSF t-tau statistically mediated the effect of CSF IGFBP-2 on atrophy in parahippocampal ($p = 0.007$), entorhinal ($p = 0.01$), inferior temporal ($p = 0.003$), and temporal pole ($p = 0.02$) regions.

Igfbp2/*IGFBP2* Is Differentially Expressed in Transgenic Mice and Selectively Expressed in Astrocytes

Igfbp2 was differentially expressed in both TASTPM (AD) and P301L tau-transgenic compared to wild-type (C57BL/6) mouse neuropathological data. In cortex, homozygote TASTPM AD mice displayed lower *Igfbp2* expression during early life and higher expression during late life compared to wild-type mice (**Figure 3A**; $F = 9.28$, $p = 0.004$). By contrast, heterozygote TASTPM AD mice showed consistently lower expression of *Igfbp2* across all ages compared to wild-type mice (**Figure 3A**; $F = 6.26$, $p = 0.016$). Cortical expression of *Igfbp2* in the P301L tau mouse model showed an expression pattern similar to TASTPM AD homozygotes, with greater expression at older ages (**Figure 3B**; $F = 5.03$, $p = 0.029$). In hippocampus, *Igfbp2* was not significantly different from wild-type expression in either TASTPM heterozygotes (**Figure 3C**; $F = 0.21$, $p = 0.21$) or homozygotes (**Figure 3C**; $F = 0.88$, $p = 0.35$). Similarly, *Igfbp2* was not significantly different in P301L tau transgenic mice compared to wild-type mice, with expression increasing over time in both genotypes (**Figure 3D**; $F = 0.19$, $p = 0.67$).

Finally, we assessed cell specificity of *IGFBP2* expression in the CNS. In both humans and mice, astrocytes expressed *IGFBP2*/*Igfbp2* most robustly (**Figure 4**). In human samples, fetal astrocytes expressed *IGFBP2* more highly than mature astrocytes (**Figure 4A**). In mice, oligodendrocyte progenitor cells and neurons were the next highest expressers of *Igfbp2* following astrocytes (**Figure 4B**).

DISCUSSION

Our study highlights several findings supporting the role of IGFBP-2 in tau-related AD pathobiology. Previous studies on AD-associated changes in CSF IGFBP-2 are conflicting; some studies demonstrate elevated levels of CSF IGFBP-2 in AD relative to age-matched controls (Salehi et al., 2008; Hertze et al., 2014), while others report similar findings to our study with no significant difference in CSF IGFBP-2 levels by diagnosis (Åberg et al., 2015; Lane et al., 2017). Furthermore, our finding that IGFBP-2 levels correlate with t-tau and p-tau in the CSF across the AD spectrum at baseline (**Table 2**) are in agreement with prior studies (Hertze et al., 2014; Åberg et al., 2015). Taken together with previous studies, our findings suggest that CSF IGFBP-2 levels associate with pathological burden across the spectrum of AD. Further experimental studies are required to elucidate a mechanistic connection between IGFBP-2 and tau pathology in AD.

A previous study reported an association between plasma IGFBP-2 levels and cross-sectional atrophy measured by SPARE-AD score in a subset of healthy controls, MCI, and AD participants from ADNI, but found no association between plasma IGFBP-2 levels and atrophy in specific brain regions (Toledo et al., 2013). More recent studies found that plasma IGFBP-2 was associated with hippocampal volumes as well as other volumetric and functional measures cross-sectionally,

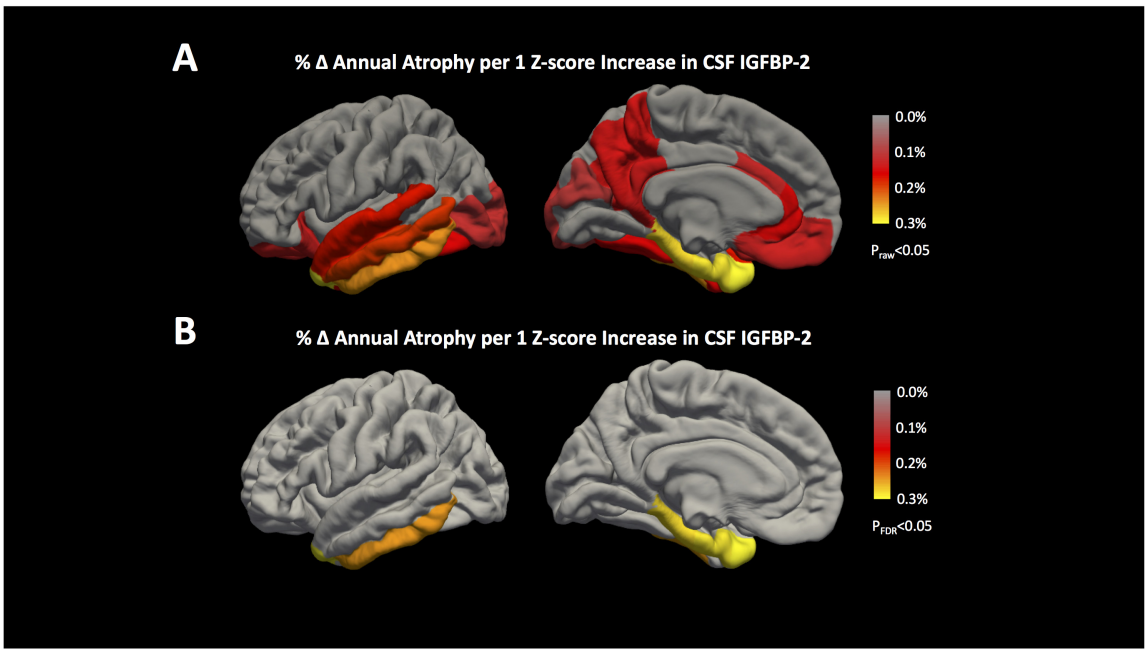

FIGURE 2 | IGFBP-2 is associated with longitudinal atrophy in entorhinal, inferior temporal, temporal pole, and parahippocampal regions. Annualized changes in atrophy rate given a 1 *z*-score increase in CSF IGFBP-2 levels are shown for all 34 cortical regions included in the Desikan Killiany Atlas (Desikan et al., 2006). The results of our analyses are shown **(A)** before correction for multiple testing ($p_{raw} < 0.05$) and **(B)** after correction for multiple testing ($p_{FDR} < 0.05$) using the FDR method (Benjamini and Hochberg, 1995). Hippocampus and amygdala are not shown, but the results of these analyses are provided in Supplementary Table 2. IGFBP-2 levels are quality controlled and transformed as previously described (Siuciak, 2011). After correction for multiple testing, greater levels of CSF IGFBP-2 are associated with higher annual rates of atrophy in entorhinal, inferior temporal, temporal pole, and parahippocampal regions.

but not longitudinally (Lane et al., 2017; Mclimans et al., 2017). A recent report demonstrated a significant relationship ($p = 0.023$) between CSF amyloid and CSF IGFBP-2, potentially conflicting with our finding (Mclimans et al., 2017). The discrepancy between these findings and ours could be explained by differences in our covariate selection. For instance, we covaried for clinical severity (CDR-SB) score rather than for baseline diagnosis and included *APOE* ε4 carrier status rather than *APOE* ε4 dosage in all analyses. We found an association between baseline CSF IGFBP-2 levels and longitudinal changes in multiple non-hippocampal brain structures (Supplementary Table 2). Surprisingly, *Igfbp2* expression in transgenic mice was significantly different from wild type mice only in cortex and not in hippocampus, which may be why we (and other groups) failed to find a robust association between IGFBP-2 and longitudinal hippocampal atrophy. Additionally, it is possible that CSF IGFBP-2 represents a more proximal measure of IGF dysregulation in the brain relative to plasma levels, which may be more variable in a limited clinical cohort.

Insulin and IGF resistance due to type 2 diabetes in human patients significantly increases the risk of developing dementia, and is associated with regional brain atrophy (Leibson et al., 1997; Last et al., 2007; Stanley et al., 2016). Furthermore, impaired brain insulin and IGF signaling induced in rats by intracerebral injection of streptozotocin results in brain atrophy and neurodegeneration (Lester-Coll et al., 2006). While the exact role of IGFBP-2 in regulating IGF signaling in the brain is unclear, evidence in mice suggests that IGFBP-2 may inhibit IGF activity (Hoeflich et al., 1999). We provide statistical evidence that the association between CSF IGFBP-2 and entorhinal, parahippocampal, inferior temporal, and temporal pole atrophy may be related to intracerebral tau (estimated using CSF tau levels).

Tau dysregulation is a hallmark of AD pathology and contributes to neuronal cell loss (Querfurth and LaFerla, 2010). As impaired IGF signaling contributes to tau dysregulation (Bedse et al., 2015), the effect of IGFBP-2 on IGF signaling may explain how IGFBP-2 contributes to tau-related brain atrophy. Similarly, previous studies in primary neurons demonstrated that IGF-I prevents amyloid-induced increases in tau phosphorylation and cell death, and IGFBP-3 was able to inhibit these effects (Watanabe et al., 2015). Although IGFBP-2 may regulate IGF signaling in neurons differently than IGFBP-3, one might speculate that IGFBP-2 binds to IGFs, blocking IGF-mediated suppression of tau phosphorylation, leading to increased levels of p-tau and promoting neuronal damage and death.

A strength of our study is the use of a thoroughly characterized cohort of healthy aging control, MCI, and AD patients, a subset of which underwent multiple MRI scans and had baseline CSF protein levels quantified. Our findings utilized multiple data types and support a role for IGFBP-2 in AD pathobiology. However, our study is limited by its observational nature, which prevents us

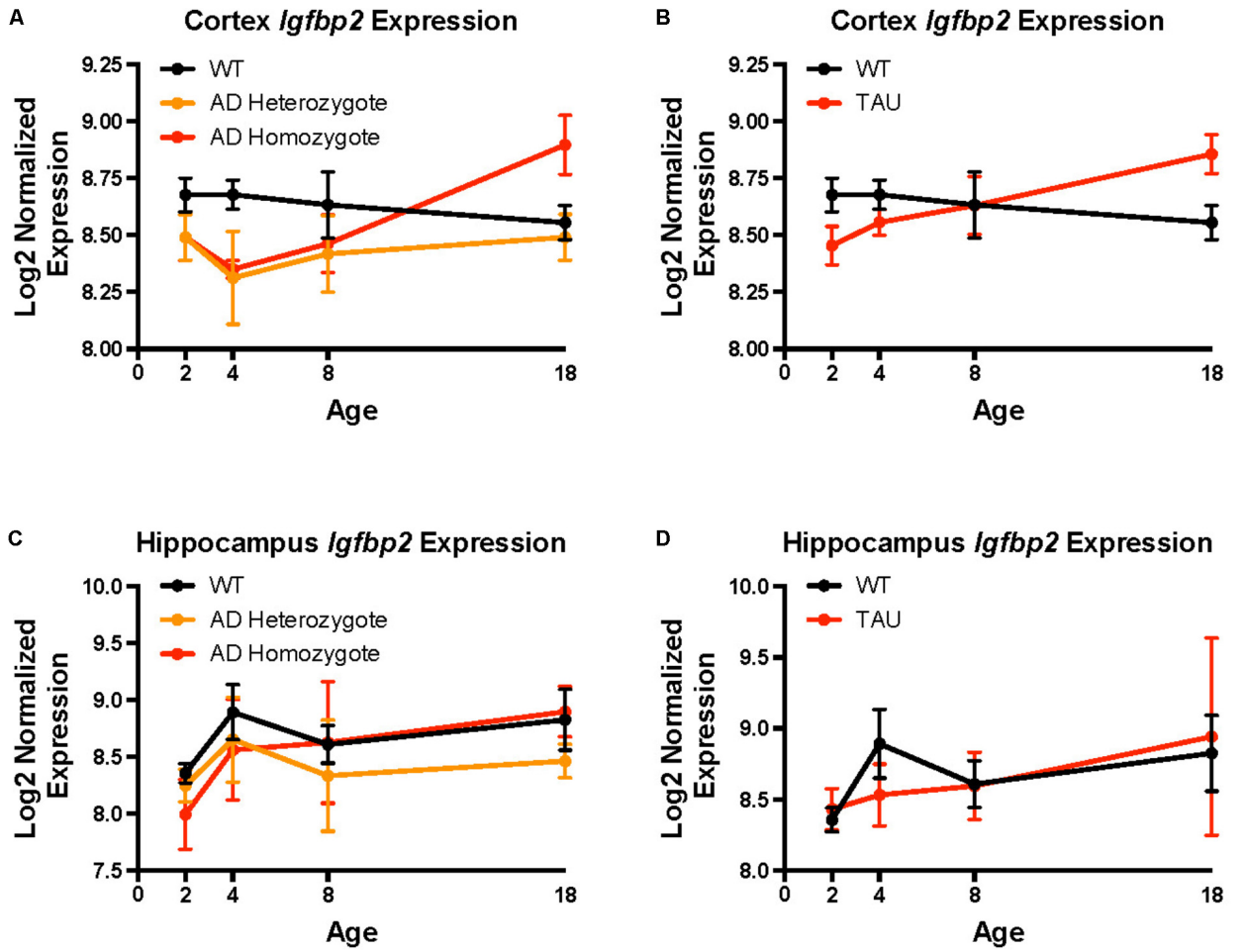

FIGURE 3 | *Igfbp2* expression levels in mouse models of neurodegeneration. Mouse *Igfbp2* expression from Mouseac (www.mouseac.org) is plotted at ages 2, 4, 8, and 18 months. **(A)** Cortex expression in the TASTPM (AD) mouse model. **(B)** Cortex expression in the P301L (TAU) transgenic mouse. **(C)** Hippocampus expression in the TASTPM (AD) mouse model. **(D)** Hippocampus expression in the P301L (TAU) transgenic mouse. The plotted points are mean expression at each age ± standard errors. Expression data was normalized using a $\log_2$ transformation and all samples were quantile normalized together. Please see Matarin et al. (2015) for additional details on experimental models and data processing.

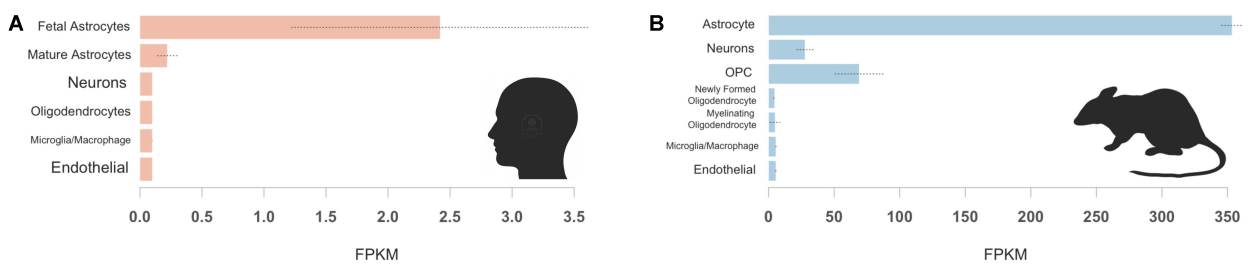

FIGURE 4 | *IGFBP2* expression by cell type in human and mouse brain samples. Human *IGFBP2* **(A)** and mouse *Igfbp2* **(B)** expression is shown for selected CNS cell types available from (http://web.stanford.edu/group/barres_lab/brainseq2/brainseq2.html). In both humans and mice, astrocytes express *IGFBP2* most robustly oligodendrocyte precursor cell (OPC).

from establishing causative relationships. Additionally, the data from murine models of neurodegenerative disease only allowed for examination of whole cerebral cortex. However, our analyses using human data highlighted parahippocampal, entorhinal, inferior temporal, and temporal pole cortex as the regions whose atrophy is most associated with CSF IGFBP-2. Thus, we cannot easily compare the neuroanatomical relationships seen in our human data with the cross-sectional mouse data. As a correlative

study, our findings suggest that CSF IGFBP-2 levels are related to AD, but do not carry any mechanistic implications. For example, although we have proposed an inhibitory role for IGFBP-2 on IGF signaling in the brain based on previous studies, others have indicated that IGFBP-2 may facilitate IGF signaling in the brain (Russo et al., 2005), and therefore elevated levels of IGFBP-2 in the brain may protect against AD pathogenesis. Thus, our results require follow-up in larger independent cohorts and experimental models to establish whether IGFBP-2 influences progression from normal cognition to AD and its potential biological role in AD pathogenesis.

In summary, we found that baseline IGFBP-2 levels correlate with t-tau and p-tau levels in the CSF of healthy aging control, MCI and AD patients. IGFBP-2 is associated with longitudinal rates of atrophy in AD-associated human cortical regions and its expression is dysregulated in transgenic mice with AD-relevant pathology. In both humans and mice, *IGFBP2/Igfbp2* is most highly expressed in astrocytes. Given the increasingly appreciated role of astrocytes in synaptic pruning during neurodegeneration (Liddelow et al., 2017), further studies may help to elucidate why this effect appears to be limited to non-hippocampal regions and how astrocyte-related metabolic disarray leads to tau pathology in AD.

AUTHOR CONTRIBUTIONS

LB conceived the study, conducted analyses, and drafted the manuscript. EG conceived the study, conducted analyses, and drafted the manuscript. NS drafted the manuscript and interpreted the data. DH processed MRI scans and provided guidance on image analysis. BM assisted with data interpretation. AD processed MRI scans, provided guidance on image analysis, and interpreted data. RD processed MRI scans, provided guidance on image as well as biomarker analyses, and drafted the manuscript. JY conceived the study, conducted analyses, and drafted the manuscript.

FUNDING

Primary support for data analyses was provided by the Larry L. Hillblom Foundation 2016-A-005-SUP (JY), NIA K01 AG049152 (JY), the Rainwater Charitable Foundation (JY), the Bluefield Project to Cure FTD (JY), the French Foundation (JY), Radiological Society of North America RMS 1741 (LB), National Alzheimer's Coordinating Center (NACC) Junior Investigator Award (RD), ASNR Foundation Alzheimer's Imaging Grant (RD), and NIA P50 AG23501 (BM). Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI; National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI was funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research and Development, LLC.; Johnson and Johnson Pharmaceutical Research and Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; NeuroTrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

ACKNOWLEDGMENTS

We thank the ADNI research participants for their longstanding dedication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2018.00476/full#supplementary-material

REFERENCES

Åberg, D., Johansson, P., Isgaard, J., Wallin, A., Johansson, J. O., Andreasson, U., et al. (2015). Increased cerebrospinal fluid level of insulin-like growth factor-II in male patients with alzheimer's disease. *J. Alzheimers Dis.* 48, 637–646. doi: 10.3233/JAD-150351

Bedse, G., Di Domenico, F., Serviddio, G., and Cassano, T. (2015). Aberrant insulin signaling in Alzheimer's disease: current knowledge. *Front. Neurosci.* 9:204. doi: 10.3389/fnins.2015.00204

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300. doi: 10.2307/2346101

Bennett, M. L., Bennett, F. C., Liddelow, S. A., Ajami, B., Zamanian, J. L., Fernhoff, N. B., et al. (2016). New tools for studying microglia in the mouse and human CNS. *Proc. Natl. Acad. Sci. U.S.A.* 113, E1738–E1746. doi: 10.1073/pnas.1525528113

Bomfim, T. R., Forny-Germano, L., Sathler, L. B., Brito-Moreira, J., Houzel, J. C., Decker, H., et al. (2012). An anti-diabetes agent protects the mouse brain from defective insulin signaling caused by Alzheimer's disease–associated Aβ oligomers. *J. Clin. Invest.* 122, 1339–1353. doi: 10.1172/JCI57256

Bonham, L. W., Desikan, R. S., and Yokoyama, J. S. (2016). The relationship between complement factor C3, APOE ε4, amyloid and tau in Alzheimer's disease. *Acta Neuropathol. Commun.* 4, 1–7. doi: 10.1186/s40478-016-0339-y

Chen, Y., Liang, Z., Blanchard, J., Dai, C. L., Sun, S., Lee, M. H., et al. (2013). A non-transgenic mouse model (icv-STZ mouse) of Alzheimer's disease: similarities to and differences from the transgenic model (3xTg-AD Mouse). *Mol. Neurobiol.* 47, 711–725. doi: 10.1007/s12035-012-8375-8375

Craig-Schapiro, R., Kuhn, M., Xiong, C., Pickering, E. H., Liu, J., Misko, T. P., et al. (2011). Multiplexed immunoassay panel identifies novel CSF biomarkers for Alzheimer's disease diagnosis and prognosis. *PLoS One* 6:e18850. doi: 10.1371/journal.pone.0018850

Deng, Y., Li, B., Liu, Y., Iqbal, K., Grundke-Iqbal, I., and Gong, C. X. (2009). Dysregulation of insulin signaling, glucose transporters, O-glcNAcylation, and phosphorylation of tau and neurofilaments in the brain: implication for Alzheimer's disease. *Am. J. Pathol.* 175, 2089–2098. doi: 10.2353/ajpath.2009.090157

Desikan, R. S., McEvoy, L. K., Thompson, W. K., Holland, D., Roddey, J. C., Blennow, K., et al. (2011). Amyloid-β associated volume loss occurs only in the presence of phospho-tau. *Ann. Neurol.* 70, 657–661. doi: 10.1002/ana.22509

Desikan, R. S., Seggonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021

Desikan, R. S., Thompson, W. K., Holland, D., Hess, C. P., Brewer, J. B., Zetterberg, H., et al. (2013). Heart fatty acid binding protein and Aβ-associated Alzheimer's neurodegeneration. *Mol. Neurodegener.* 8:39. doi: 10.1186/1750-1326-8-39

Desikan, R. S., Thompson, W. K., Holland, D., Hess, C. P., Brewer, J. B., Zetterberg, H., et al. (2014). The role of clusterin in amyloid-β-associated neurodegeneration. *JAMA Neurol.* 71, 180–187. doi: 10.1001/jamaneurol.2013.4560

Fennema-Notestine, C., Gamst, A. C., Quinn, B. T., Pacheco, J., Jernigan, T. L., Thal, L., et al. (2007). Feasibility of multi-site clinical structural neuroimaging studies of aging using legacy data. *Neuroinformatics* 5, 235–245. doi: 10.1007/s12021-007-9003-9009

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., et al. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33, 341–355. doi: 10.1016/S0896-6273(02)00569-X

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6

Freiherr, J., Hallschmid, M., Frey, W. H., Brünner, Y. F., Chapman, C. D., Hölscher, C., et al. (2013). Intranasal insulin as a treatment for Alzheimer's disease: a review of basic research and clinical evidence. *CNS Drugs* 27, 505–514. doi: 10.1007/s40263-013-0076-78

Frölich, L., Blum-Degen, D., Bernstein, H. G., Engelsberger, S., Humrich, J., Laufer, S., et al. (1998). Brain insulin and insulin receptors in aging and sporadic Alzheimer's disease. *J. Neural Transm.* 105, 423–438. doi: 10.1007/s007020050068

Hertze, J., Nägga, K., Minthon, L., and Hansson, O. (2014). Changes in cerebrospinal fluid and blood plasma levels of IGF-II and its binding proteins in Alzheimer's disease: an observational study. *BMC Neurol.* 14:64. doi: 10.1186/1471-2377-14-64

Hoeflich, A., Wu, M., Mohan, S., Föll, J., Wanke, R., Froehlich, T., et al. (1999). Overexpression of insulin-like growth factor-binding protein-2 in transgenic mice reduces postnatal body weight gain. *Endocrinology* 140, 5488–5496. doi: 10.1210/endo.140.12.7169.

Holland, D., and Dale, A. M. (2011). Nonlinear registration of longitudinal images and measurement of change in regions of interest. *Med. Image Anal.* 15, 489–497. doi: 10.1016/j.media.2011.02.005

Holland, D., McEvoy, L. K., and Dale, A. M. (2012). Unbiased comparison of sample size estimates from longitudinal structural measures in ADNI. *Hum. Brain Mapp.* 33, 2586–2602. doi: 10.1002/hbm.21386

Holly, J., and Perks, C. (2006). The role of insulin-like growth factor binding proteins. *Neuroendocrinology* 83, 154–160. doi: 10.1159/000095523

Holscher, C., and Li, L. (2010). New roles for insulin-like hormones in neuronal signalling and protection: new hopes for novel treatments of Alzheimer's disease? *Neurobiol. Aging* 31, 1495–1502. doi: 10.1016/j.neurobiolaging.2008.08.023

Kang, J. H., Vanderstichele, H., Trojanowski, J. Q., and Shaw, L. M. (2012). Simultaneous analysis of cerebrospinal fluid biomarkers using microsphere-based xMAP multiplex technology for early detection of Alzheimer's disease. *Methods* 56, 484–493. doi: 10.1016/j.ymeth.2012.03.023

Lane, E. M., Hohman, T. J., and Jefferson, A. L. (2017). Insulin-like growth factor binding protein-2 interactions with Alzheimer's disease biomarkers. *Brain Imaging Behav.* 11, 1779–1786. doi: 10.1007/s11682-016-9636-9630

Last, D., Alsop, D. C., Abduljalil, A. M., Marquis, R. P., de Bazelaire, C., Hu, K., et al. (2007). Global and regional effects of type 2 diabetes on brain tissue volumes and cerebral vasoreactivity. *Diabetes Care* 30, 1193–1199. doi: 10.2337/dc06-2052

Leibson, C. L., Rocca, W. A., Hanson, V. A., Cha, R., Kokmen, E., O'Brien, P. C., et al. (1997). Risk of dementia among persons with diabetes mellitus: a population-based cohort study. *Am. J. Epidemiol.* 145, 301–308. doi: 10.1093/oxfordjournals.aje.a009106

Lester-Coll, N., Rivera, E. J., Soscia, S. J., Doiron, K., Wands, J. R., and de la Monte, S. M. (2006). Intracerebral streptozotocin model of type 3 diabetes: relevance to sporadic Alzheimer's disease. *J. Alzheimers Dis.* 9, 13–33. doi: 10.3233/JAD-2006-9102

Liddelow, S. A., Guttenplan, K. A., Clarke, L. E., Bennett, F. C., Bohlen, C. J., Schirmer, L., et al. (2017). Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* 541, 481–487. doi: 10.1038/nature21029

de la Monte, S. M. (2012). Brain insulin resistance and deficiency as therapeutic targets in Alzheimer's disease. *Curr. Alzheimer Res.* 9, 35–66. doi: 10.2174/156720512799015037

Matarin, M., Salih, D. A., Yasvoina, M., Cummings, D. M., Guelfi, S., Liu, W., et al. (2015). A genome-wide gene-expression analysis and database in transgenic mice during development of amyloid or tau pathology. *Cell Rep.* 10, 633–644. doi: 10.1016/j.celrep.2014.12.041

Matthews, B. R. (2010). Alzheimer disease update. *Continuum (Minneap. Minn)* 16, 15–30. doi: 10.1212/01.CON.0000368210.41093.4e

Mattson, M. P. (1997). Cellular actions of beta-amyloid precursor protein and its soluble and fibrillogenic derivatives. *Physiol. Rev.* 77, 1081–1132. doi: 10.1152/physrev.1997.77.4.1081

Mcevoy, L. K., Fennema-notestine, C., Roddey, J. C., Holland, D., Pung, C. J., Brewer, J. B., et al. (2009). Alzheimer disease: quantitative structural neuroimaging for detection and prediction of clinical and purpose: methods: results?: conclusion?. *Radiology* 251, 195–205. doi: 10.1148/radiol.2511080924

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology* 34, 939–944. doi: 10.1212/WNL.34.7.939

Mclimans, K. E., Webb, J. L., Anantharam, V., Kanthasamy, A., Willette, A. A., Nutrition, H., et al. (2017). Peripheral versus central index of metabolic dysfunction and associations with clinical and pathological outcomes in Alzheimer's disease. *J. Alzheimers Dis.* 60, 1313–1324.

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 43, 2412–2414. doi: 10.1212/WNL.43.11.2412-a

Perl, D. P. (2010). Neuropathology of Alzheimer's disease. *Mt. Sinai J. Med.* 77, 32–42. doi: 10.1002/msj.20157

Preacher, K. J., and Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behav. Res. Methods Instrum. Comput.* 36, 717–731. doi: 10.3758/BF03206553

Querfurth, H. W., and LaFerla, F. M. (2010). Alzheimer's disease. *N. Engl. J. Med.* 362, 329–344. doi: 10.1056/NEJMra0909142

Reger, M. A., Watson, G. S., Green, P. S., Wilkinson, C. W., Baker, L. D., Cholerton, B., et al. (2008). Intranasal insulin improves cognition and modulates beta-amyloid in early AD. *Neurology* 70, 440–448. doi: 10.1212/01.wnl.0000265401.62434.36

Rivera, E. J., Goldin, A., Fulmer, N., Tavares, R., Wands, J. R., and de la Monte, S. M. (2005). Insulin and insulin-like growth factor expression and function deteriorate with progression of Alzheimer's disease: link to brain reductions in acetylcholine. *J. Alzheimers Dis.* 8, 247–268. doi: 10.3233/JAD-2005-8304

Russo, V. C., Gluckman, P. D., Feldman, E. L., and Werther, G. A. (2005). The insulin-like growth factor system and its pleiotropic functions in brain. *Endocr. Rev.* 26, 916–943. doi: 10.1210/er.2004-2024

Salehi, Z., Mashayekhi, F., and Naji, M. (2008). Insulin like growth factor-1 and insulin like growth factor binding proteins in the cerebrospinal fluid and serum from patients with Alzheimer's disease. *Biofactors* 33, 99–106. doi: 10.1002/biof.5520330202

Salkovic-Petrisic, M., and Hoyer, S. (2007). "Central insulin resistance as a trigger for sporadic Alzheimer-like pathology: an experimental approach," in *Neuropsychiatric Disorders An Integrative Approach*, eds M. Gerlach, J. Deckert, K. Double, and E. Koutsilieri (Vienna: Springer), 217–233. doi: 10.1007/978-3-211-73574-9-28

Shaw, L. M., Vanderstichele, H., Knapik-Czajka, M., Clark, C. M., Aisen, P. S., Petersen, R. C., et al. (2009). Cerebrospinal fluid biomarker signature in alzheimer's disease neuroimaging initiative subjects. *Ann. Neurol.* 65, 403–413. doi: 10.1002/ana.21610

Siuciak, J. (2011). *Biomarkers Consortium Data Primer*. Available at: https://adni.loni.usc.edu/wp-content/uploads/2012/01/2011Dec28-Biomarkers-Consortium-Data-Primer-FINAL1.pdf

Stanley, M., Macauley, S. L., and Holtzman, D. M. (2016). Changes in insulin and insulin signaling in Alzheimer's disease: cause or consequence? *J. Exp. Med.* 213, 1375–1385. doi: 10.1084/jem.20160493

Steen, E., Terry, B. M., Rivera, E. J., Cannon, J. L., Neely, T. R., Tavares, R., et al. (2005). Impaired insulin and insulin-like growth factor expression and signaling mechanisms in Alzheimer's disease–is this type 3 diabetes? *J. Alzheimers Dis.* 7, 63–80. doi: 10.3233/JAD-2005-7107

Stockhorst, U., de Fries, D., Steingrueber, H. J., and Scherbaum, W. A. (2004). Insulin and the CNS: effects on food intake, memory, and endocrine parameters and the role of intranasal insulin administration in humans. *Physiol. Behav.* 83, 47–54. doi: 10.1016/j.physbeh.2004.07.022

van Dam, P. S., and Aleman, A. (2004). Insulin-like growth factor-I, cognition and brain aging. *Eur. J. Pharmacol.* 490, 87–95. doi: 10.1016/j.ejphar.2004.02.047

Takeda, S., Sato, N., Uchio-Yamada, K., Sawada, K., Kunieda, T., Takeuchi, D., et al. (2010). Diabetes-accelerated memory dysfunction via cerebrovascular inflammation and Aβ deposition in an Alzheimer mouse model with diabetes. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7036–7041. doi: 10.1073/pnas.1000645107

Talbot, K., Wang, H. Y., Kazi, H., Han, L. Y., Bakshi, K. P., Stucky, A., et al. (2012). Demonstrated brain insulin resistance in Alzheimer's disease patients is associated with IGF-1 resistance, IRS-1 dysregulation, and cognitive decline. *J. Clin. Invest.* 122, 1316–1338. doi: 10.1172/JCI59903

Tham, A., Nordberg, A., Grissom, F. E., Carlsson-Skwirut, C., Viitanen, M., and Sara, V. R. (1993). Insulin-like growth factors and insulin-like growth factor binding proteins in cerebrospinal fluid and serum of patients with dementia of the Alzheimer type. *J. Neural Transm. Park. Dis. Dement. Sect.* 5, 165–176. doi: 10.1007/BF02257671

Toledo, J. B., Da, X., Bhatt, P., Wolk, D. A., Arnold, S. E., Shaw, L. M., et al. (2013). Relationship between plasma analytes and SPARE-AD defined brain atrophy patterns in ADNI. *PLoS One* 8:e55531. doi: 10.1371/journal.pone.0055531

Vardy, E. R., Rice, P. J., Bowie, P. C., Holmes, J. D., Grant, P. J., and Hooper, N. M. (2007). Increased circulating insulin-like growth factor-1 in late-onset Alzheimer's disease. *J. Alzheimers Dis.* 12, 285–90. doi: 10.3233/JAD-2007-12401

Watanabe, K., Uemura, K., Asada, M., Maesako, M., Akiyama, H., Shimohama, S., et al. (2015). The participation of insulin-like growth factor-binding protein 3 released by astrocytes in the pathology of Alzheimer's disease. *Mol. Brain* 8:82. doi: 10.1186/s13041-015-0174-172

Watson, G. S., and Craft, S. (2003). The role of insulin resistance in the pathogenesis of Alzheimer's disease. *CNS Drugs* 17, 27–45. doi: 10.2165/00023210-200317010-200317013

Wechsler, D. (1987). *Wechsler Memory Scale—Revised Manual*. San Antonio, TX: Psychological Corporation.

Zhang, Y., Chen, K., Sloan, S. A., Bennett, M. L., Scholze, A. R., O'Keeffe, S., et al. (2014). An rna-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* 34, 11929–11947. doi: 10.1523/JNEUROSCI.1860-14.2014

Zhang, Y., Sloan, S. A., Clarke, L. E., Caneda, C., Plaza, C. A., Blumenthal, P. D., et al. (2016). Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse. *Neuron* 89, 37–53. doi: 10.1016/j.neuron.2015.11.013

Zhao, L., Teter, B., Morihara, T., Lim, G. P., Ambegaokar, S. S., Ubeda, O. J., et al. (2004). Insulin-degrading enzyme as a downstream target of insulin receptor signaling cascade: implications for Alzheimer's disease intervention. *J. Neurosci.* 24, 11120–11126. doi: 10.1523/JNEUROSCI.2860-04.2004

Conflict of Interest Statement: RD is an editor for this special issue of Frontiers in Neuroscience.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 27 August 2018
doi: 10.3389/fnins.2018.00592

Targeted Sequencing of Alzheimer Disease Genes in African Americans Implicates Novel Risk Variants

*Mark W. Logue[1,2,3,4], Daniel Lancour[3], John Farrell[3], Irina Simkina[3], M. Daniele Fallin[5], Kathryn L. Lunetta[4] and Lindsay A. Farrer[3,4,6,7]**

[1] National Center for Posttraumatic Stress Disorder (PTSD), United States Department of Veterans Affairs, Boston Healthcare System, Boston, MA, United States, [2] Department of Psychiatry, Boston University School of Medicine, Boston University, Boston, MA, United States, [3] Biomedical Genetics, Department of Medicine, Boston University School of Medicine, Boston University, Boston, MA, United States, [4] Department of Biostatistics, Boston University School of Public Health, Boston University, Boston, MA, United States, [5] Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States, [6] Departments of Neurology and Ophthalmology, Boston University School of Medicine, Boston University, Boston, MA, United States, [7] Department of Epidemiology, Boston University School of Public Health, Boston, MA, United States

The genetic architecture of late-onset Alzheimer disease (AD) in African Americans (AAs) differs from that in persons of European ancestry. In addition to *APOE*, genome-wide association studies (GWASs) of AD in AA samples have implicated *ABCA7*, *COBL*, and *SLC10A2* as AA-AD risk genes. Previously, we identified by whole exome sequencing a small number of AA AD cases and subsequent genotyping in a large AA sample of AD cases and controls association of AD risk with a pair of rare missense variants in *AKAP9*. In this study, we performed targeted deep sequencing (including both introns and exons) of approximately 100 genes previously linked to AD or AD-related traits in an AA cohort of 489 AD cases and 472 controls to find novel AD risk variants. We observed association with an 11 base-pair frame-shift loss-of-function (LOF) variant in *ABCA7* (rs567222111) for which the evidence was bolstered when combined with data from a replication AA cohort of 484 cases and 484 controls (*OR* = 2.42, p = 0.022). We also found association of AD with a rare 9 bp deletion (rs371245265) located very close to the *AKAP9* transcription start site (rs371245265, *OR* = 10.75, p = 0.0053). The most significant findings were obtained with a rare protective variant in *F5* (*OR* = 0.053, $p = 6.40 \times 10^{-5}$), a gene that was previously associated with a brain MRI measure of hippocampal atrophy, and two common variants in *KIAA0196* (*OR* = 1.51, $p < 8.6 \times 10^{-5}$). Gene-based tests of aggregated rare variants yielded several nominally significant associations with *KANSL1*, *CNN2*, and *TRIM35*. Although no associations passed multiple test correction, our study adds to a body of literature demonstrating the utility of examining sequence data from multiple ethnic populations for discovery of new and impactful risk variants. Larger sample sizes will be needed to generate well-powered epidemiological investigations of rare variation, and functional studies are essential for establishing the pathogenicity of variants identified by sequencing.

Keywords: Alzheimer disease, sequencing, *AKAP9*, *ABCA7*, rare variant, African Americans

OPEN ACCESS

Edited by:
Celeste Karch, Washington University in St. Louis, United States

Reviewed by:
Sheng Chih Jin, Yale University, United States
Manav Kapoor, Icahn School of Medicine at Mount Sinai, United States

***Correspondence:**
Lindsay A. Farrer
farrer@bu.edu

Specialty section:
This article was submitted to Neurogenomics, a section of the journal Frontiers in Neuroscience

Received: *05 March 2018*
Accepted: *07 August 2018*
Published: *27 August 2018*

Citation:
Logue MW, Lancour D, Farrell J, Simkina I, Fallin MD, Lunetta KL and Farrer LA (2018) Targeted Sequencing of Alzheimer Disease Genes in African Americans Implicates Novel Risk Variants. Front. Neurosci. 12:592. doi: 10.3389/fnins.2018.00592

INTRODUCTION

Studies of common genetic variants have identified many gene loci that influence risk of late-onset Alzheimer disease (AD) in persons of European ancestry (EA), most notably the *APOE* ε2 and ε4 alleles which confer strong protective and deleterious effects, respectively (Saunders et al., 1993; Corder et al., 1994), as well as more than 20 modest effect loci (odds ratios between 1.1 and 1.3) including *BIN1, CR1, ABCA7*, *CLU*, *PICALM*, and the *MS4A* gene region (Lambert et al., 2013). Extensions of these findings and the contributions of additional loci have emerged from investigations of non-EA cohorts, African Americans (AAs) in particular (Reitz et al., 2013a; Mez et al., 2017). The risk of AD is greater in AAs than EAs, however, paradoxically, ε4 has a weaker effect in AAs than EAs (Farrer et al., 1997; Reitz et al., 2013a). These observations and greater genetic diversity among persons with African ancestry suggest that the genetic architecture for AD includes some variants and loci that differ from EAs. Several genome-wide association studies (GWAS) in AAs (Logue et al., 2011; Kamboh et al., 2012; Reitz et al., 2013a) confirmed the role of several genes identified initially in EAs, most notably *APOE* and *ABCA7*. The association peak in *ABCA7* is ascribed to different SNPs in EAs (rs4147929) and AAs (rs115550680; Lambert et al., 2013; Reitz et al., 2013a). Gene resequencing studies have revealed multiple rare *ABCA7* deletions causing missense loss-of-function (LOF) mutations in EAs (Cukier et al., 2016; N'Songo et al., 2017). Cukier et al. (2016) identified a 44 base pair (bp) frameshift deletion in *ABCA7* (rs142076058) that is in linkage disequilibrium (LD) with rs115550680 and thus may be the functional variant underlying the observed association. A recent exome sequencing investigation in an AA cohort of 198 AD cases and 304 controls examined 20 putative AD risk genes implicated by GWAS in EAs, and found nominally (uncorrected) significant associations with two *ABCA7* variants (rs3764647 and rs3752239) and with gene-based tests of coding variants in *MS4A6A*, *PTK2B*, and *ZCWPW1* (N'Songo et al., 2017).

Novel AD loci have been identified in other studies of AA samples. Mez et al. (2017) identified GWAS significant associations with SNPs in *COBL (rs112404845)* and *SLC10A2 (rs16961023)* in a GWAS using an informed conditioning approach. A WES study of seven AA cases followed by genotyping using a staged design in AA cohorts containing 422 cases and 394 controls (stage 1) and 1,037 cases and 1,869 controls (stage 2) identified association with two rare AA-specific highly correlated variants in *AKAP9,* rs144662445 (*OR* = 2.75) and rs149979685 (*OR* = 3.61) (Logue et al., 2014).

These studies confirm the utility of examining African-descent samples to identify new AD risk variants in known AD genes as well as novel AD loci. In this study, we performed targeted sequencing in a discovery cohort containing approximately 1,000 AAs to identify new potentially causal variants in risk genes previously implicated in AD risk in AAs (*ABCA7*, *AKAP9*, *COBL*, *MS4A6A*, *PTK2B*, *SLC10A2*, and *ZCWPW1*) or in AD and related traits in other populations.

MATERIALS AND METHODS

Samples

The targeted gene sequencing sample included AA subjects primarily from two cohorts: the Multi-Institutional Research on Alzheimer Genetic Epidemiology (MIRAGE, 113 AD cases, 131 controls) Study (Green et al., 2002) and the Genetic and Environmental Risk Factors for Alzheimer Disease Among African Americans (GenerAAtions, 222 AD cases, 190 controls) Study (Logue et al., 2011). MIRAGE is a family-based study of clinic-based AD cases and their first-degree relatives. The GenerAAtions study includes unrelated individuals ascertained through the Henry Ford Health System. In addition, we obtained DNA samples and phenotypic data from the National Cell Repository for Alzheimer Disease (NCRAD) that were aggregated from the Ibadan/Indianapolis (INDY) study (Hendrie et al., 1995; Sahota et al., 1997; Gureje et al., 2006), the African American Alzheimer's Disease Genetics (AAG) study (Meier et al., 2012), the National Institute on Aging Alzheimer's Disease Centers (ADC) (Jun et al., 2010), and the National Institute on Aging Late-Onset Alzheimer's Disease (NIA-LOAD) Family Study (Lee et al., 2008). The Indianapolis/Ibadan study comprises elderly AA residents from Indianapolis (community dwelling and nursing home residents) and African-descent residents of Ibadan, Nigeria. The AAG study and ADC cohort include cases and controls ascertained at more than 30 sites across the United States. The NIA-LOAD Study includes families with multiple AD cases and unaffected members and an independent set of cognitively screened controls.

The discovery cohort included 489 cases and 472 controls from the MIRAGE and GenerAAtions studies supplemented with 154 cases and 151 controls from the AAG and Ibadan studies. The replication cohort consisted of additional samples from the AAG, ADC, Indy/Ibadan, and NIA-LOAD studies (484 AD cases, 484 controls). Characteristics of the discovery and replication cohorts are presented in **Table 1**. Further details about subject ascertainment and classification, including genetic screening for ancestry mismatches, were reported elsewhere (Reitz et al., 2013a). The diagnosis of AD in all cohorts was made according to established criteria (McKhann et al., 2011) and all controls were screened to be cognitively normal.

Sequencing Methods

The samples in the discovery cohort were sent to the McDonnell Genome Institute at Washington University[1] for targeted sequencing. The targeted regions included genes previously associated with AD in AAs (*ABCA7*, *AKAP9*, *COBL*, *MS4A6A*, *PTK2B*, *SLC10A2*, and *ZCWPW1*) and approximately 100 other provisional and confirmed genes and regions that were identified by candidate gene and GWAS approaches in studies of AD and AD-related traits (Saunders et al., 1993; Farrer et al., 2000; Meng et al., 2006; Rogaeva et al., 2007; Vardarajan et al., 2012; Lambert et al., 2013; Reitz et al., 2013b; Jun et al., 2014; Logue et al., 2014; Wetzel-Smith et al., 2014; Jun et al., 2016; Chung et al., 2017; Mez et al., 2017) (**Supplementary Table S1**). Nimblegen probes

[1] http://genome.wustl.edu

TABLE 1 | Sample size and demographics for discovery and replication cohorts.

Discovery Data	Site	Sample size	Cases			Controls		
			N (%)	*N* male (%)	mean age at onset (*SD*)	*N* (%)	*N* male (%)	mean age at exam (*SD*)
	MIRAGE	244	113 (46.31%)	27 (23.89)	71.14 (9.25)	131 (53.69%)	40 (30.53)	69.77 (10.16)
	GenerAAtions	412	222 (46.12%)	97 (43.69)	77.27 (6.64)	190 (53.88%)	77 (40.53)	78.38 (6.51)
	Ibadan	119	60 (50.42)	8 (13.33)	84.62 (7.37)	59 (49.58)	29 (49.15)	94.4* (3.2)
	AAG	186	94 (50.54)	26 (27.66)	80.44 (5.04)	92 (49.46)	14 (15.22)	79.89 (3.34)
	Total discovery	961	489 (50.83)			472 (49.17)		
Replication Data	**Cohort**	**Sample size**	**Cases**			**Controls**		
			N (%)	*n* male (%)	mean age at onset (*SD*)	*N* (%)	*N* male (%)	mean age at exam (*SD*)
	AAG	183	49 (26.7)	13 (26.53)	68.08 (3.81)	134 (73.22)	28 (20.9)	72.52 (2.35)
	ADC	89	73 (82.02)	26 (35.62)	76.22 (6.92)	16 (17.98)	5 (31.25)	77.31 (7.46)
	Ibadan	38	19 (50.00)	3 (15.79)	78.48 (6.30)	19 (50.00)	6 (31.58)	91.45* (4.23)
	INDY	354	171 (48.31)	58 (33.92)	84.44 (6.53)	183 (51.69)	38 (20.77)	93.81* (2.9)
	NIALOAD	304	172 (56.58)	61 (35.47)	77.11 (7.83)	132 (43.42)	45 (34.09)	72.6 (7.82)
	Total replication	968	484 (50.00)			484 (50.00)		

**Controls selected to have a high mean age.*

(Roche Nimblegen, Madison, WI, United States) were generated to cover all non-repetitive exonic, intronic, and intergenic sequence and 5,000 bp upstream and 1,000 bp downstream of gene boundaries including all isoforms totaling approximately nine Mb of genomic sequence. Only exons were targeted for *SLC10A2* and *COBL* because these associations (Mez et al., 2017) were not known at the time the capture design was proposed and the limited amount of genomic sequence that could be added to the capture at this stage. The capture design included 10,906 capture targets and had 92.7% estimated coverage of 122 targeted regions, with gaps due to repetitive sequence.

Samples were assessed for volume and concentration by the Genome Center using either a Qubit or a VarioSkan assay prior to sequencing. All but eight had >250 ng starting material. The sequencing was done in two waves. The first wave included 667 samples from the MIRAGE and GenerAAtions cohorts. Libraries were captured in sets of 66 and 67 samples per pool and each pool was run in two lanes of an Illumina Hiseq2500 1T platform. The remaining discovery cohort samples were sequenced in the second wave using the same capture probes in pools of 90 samples each, and each sample was run on 2 lanes of an Illumina HiSeq4000 platform. Valid sequence data were available for a discovery cohort including 489 cases and 472 controls. The median number of reads per sample was 14,322,643 (range 6,175,120–25,567,585). The median number of reads was greater for the samples run on the HiSeq4000 platform (median reads/sample for batch 1 = 12,809,719, median reads/sample for batch 2 = 16,804,253). In batch 1, the number of reads/sample for the MIRAGE Study samples (median = 12,331,728) was significantly less than for the GenerAAtions samples (median = 17,217,602, $P = 3.51 \times 10^{-5}$). The number of reads per sample for the second batch of sequencing did not vary by cohort ($p = 0.17$). Importantly, the number of reads per sample was not associated with AD status in either batch or in the combined discovery sample (all $p > 0.3$). Across samples, the median percentage of bases with more than 10 reads was 94.60 and the mean coverage depth was 155.7.

Sanger Sequencing

Genotyping for the *ABCA7* deletion polymorphism rs567222111 was performed in the replication sample by GENEWIZ (GENEWIZ LLC, South Plainfield NJ, United States[2]) using bi-directional Sanger sequencing. Sequencing was repeated for samples that did not yield a reliable genotype call in the first run. Validity of the Sanger sequencing assay was demonstrated by verifying genotype calls for 10 samples which had been identified as having the deletion by targeted-sequencing.

Data Processing and Quality Control

The 126 bp paired-end reads were aligned to the GRCh37 +Decoy reference with BWA MEM version 0.7.10-r789. Variant genotypes were jointly called within the targeted regions using the GATK 3.7 pipeline. The "best practices" pipeline included steps for duplicate removal, local realignment near indels, base quality score recalibration, and variant quality score recalibration. GATK yielded calls for 230,595 variants. Annotation of the variants was performed with SnpEff and SnpSift version 4.3i (Cingolani et al., 2012). According to SnpEff, these variants mapped to 151 protein-coding genes. Variants that were not assigned a "PASS" rating by GATK (n = 11,808) were excluded from association analyses. We also excluded variants in the HLA region (n = 24,297) due to difficulties in mapping the repetitive sequence and variants in the *APOE* region (n = 197) due to difficulties discerning associations in this region that are independent

[2]http://www.genewiz.com

of *APOE* (Jun et al., 2017). However, we did use sequence calls to derive *APOE* isoform genotypes for QC purposes (see description below). Another 5,147 variants occurring only in subjects with missing phenotype information were excluded. After these filtering steps, 189,145 variants remained. From this point forward, the pipeline differed for single variant association tests and the gene based tests. For the single variant test, variants observed only once (n = 66,278) were excluded. Genotypes with quality scores <30 were set to missing and variants with a missing rate of >20% were excluded (n = 18,526). After these filtering steps, there remained 104,341 variants for analyses. For the gene based tests, we included singleton variants but excluded variants with a mapping quality of less than 30 (n = 3,748 of 189,145). We excluded variants with minor allele frequency (MAF) in the discovery cohort >5% (32,935). One hundred seventy-three of these variants labeled as "High Impact" according to SNPeff (includes LOF variants and deletions) and 1,079 missense SNPs predicted to be possibly or probably damaging according to Polyphen2 (Adzhubei et al., 2010) were included in the gene-based analyses.

As a quality control check, we compared genotypes for *APOE* and two rare *AKAP9* missense variants (rs144662445 and rs149979685) in MIRAGE cohort subjects that were generated previously by direct genotyping to those derived from targeted sequencing. The two methods agreed for 236 of 237 *APOE* genotype calls. Among 190 subjects with overlapping genotype and sequencing data for the *AKAP9* variants, rare variant calls in three individuals (each with both variants) were concordant.

Statistical Analysis

We applied a hypothesis-driven four-stage design which prioritized variants most likely to have high impact on transcript structure or function in order to minimize the penalty associated with performing more than 100,000 tests. Specifically, variants were selected for analysis if they were (1) predicted to result in loss of function according to the SNPeff annotation, which includes nonsense (stop site) and splice site variants, out of frame deletions/insertions, and large exon-removing deletions (MacArthur et al., 2012), (2) predicted to be a missense variant according to SNPeff, and (3) within 50 base pairs (bp) of transcription start sites (position determined via the Eukaryotic Promotor Database[3]). We then examined (4) all variants (intronic and exonic) regardless of potential impact. To avoid model instability that can occur with logistic or GEE or mixed models when applied to rare variants, association of AD with individual variants was evaluated using a X^2 case:control allele test without continuity correction as implemented in PLINK v1.9 (Chang et al., 2015). For particular variants of interest identified in the allele test, we additionally checked for bias due to relatedness within the MIRAGE cohort as well as potential effects due to population substructure by computing a WALD test using a logistic mixed model in the R GMMAT package (Chen et al., 2016) including as covariates the first three principle components (PCs) for ancestry. The GMMAT package incorporates information from the relationship matrix which we computed from the genetic data in PLINK v1.9 based on 4,569 common (MAF>5%) variants from the sequence data remaining after trimming for LD (plink filter –indep-pairwise 5 20.04). PCs were also computed based on common LD-trimmed SNPs using PLINK. Gene based tests were performed for the 151 protein coding genes (as identified by the SNPeff annotation) using the variable threshold burden test (Price et al., 2010) and the collapsing burden test methods (Li and Leal, 2008) implemented in EPACTS[4] which incorporates information about related subjects in the sample. The correlation matrix for related subjects for the gene based test was estimated from the sequence data. LD estimates for 1000 genomes data were obtained using LDlink[5]. LD estimates for the sequencing results from the AD cohort were estimated using PLINK v1.9 with the –rsq dprime option.

This study, involving use of repository data and biospecimens, was approved by the Boston University Institutional Review Board.

RESULTS

Genes Previously Associated With AD in AAs

LOF Variants

In the seven genes previously associated with AD in AAs, nominally significant associations were observed for a rare LOF variant in *MS4A6A* observed only in controls (rs140130948, p = 0.013) and an 11 bp *ABCA7* deletion (rs567222111, OR = 3.57, p = 0.038) which had an estimated allele frequency (AF) of 1.1% in cases and 0.32% in controls (**Table 2**). This association remained significant in the mixed model adjusting for relatedness within the sample and including PCs for ancestry (OR = 3.65, p = 0.049). This deletion had a stronger impact on AD risk than the more common 44 bp *ABCA7* deletion (rs142076058) which was previously reported to be associated with AD in an AA cohort (OR = 1.81) (Cukier et al., 2016) but not in our sample (OR = 1.27, p = 0.16). The evidence for association with rs567222111 in the replication sample was not significant, but had the same effect direction (OR for the deletion = 1.84, p = 0.22), and the significance in the combined discovery and replication samples was greater than in the discovery sample alone (OR = 2.42, p = 0.022). No LOF variants were observed in *AKAP9*, *COBL*, *PTK2B*, *SLC10A2, or ZCWPW1*.

Missense Variants

Association tests were nominally significant for 14 of 172 missense variants tested in the seven genes previously associated with AD in AAs including a common SNP in *ABCA7* (rs5985184, p = 0.0043) and the rare missense variants in *AKAP9*, rs149979685 (OR = 10.73, p = 0.0046) and rs144662445 (OR = 6.35, p = 0.0054), previously identified in a sample that overlaps substantially with the discovery cohort in this

[3]http://epd.vital-it.ch/index.php

[4]http://genome.sph.umich.edu/wiki/EPACTS

[5]https://analysistools.nci.nih.gov/LDlink/

TABLE 2 | Association between AD and loss of function (LOF) variants observed in African American AD genes.

Gene	Ch.	BP	rsID	Effect Allele	% AFR	% Cases	% Ctrls	Alt. Allele	*OR*	*P*
MS4A6A	11	59,939,727	rs140130948	G	0.15	0.00	0.64	C	0.00	0.013
ABCA7	19	1,044,707	rs567222111	G	0.83	1.13	0.32	GGGGCACCTGGT	3.57	0.038
ABCA7	19	1,041,352	rs3752229	G	0.15	0.41	0.95	A	0.43	0.15
ABCA7	19	1,056,244	rs113809142	G	0.00	0.00	0.21	T	0.00	0.15
ABCA7	19	1,046,906	rs142076058	G	6.73	9.20	7.42	GCTGCGGGACAC CATGCGCGCCAT GGGGCTCAGCC GCGCGGTGCT	1.27	0.16
ABCA7	19	1,058,727	rs556286113	T	0.15	0.20	0.00	C	NA	0.16
ABCA7	19	1,043,395	rs77403558	T	0.30	0.31	0.11	A	2.90	0.33
MS4A6A	11	59,946,302	rs598862	C	30.18	25.87	26.69	T	0.96	0.68
MS4A6A	11	59,940,500	rs138650483	T	0.00	0.10	0.11	C	0.97	0.98

Effect allele represents the minor allele; % AFR represents the estimated effect allele frequency in the 1000 Genomes African cohort; % Cases represents the estimated effect allele frequency in AD cases; % Controls represents the estimated effect allele frequency in controls.

TABLE 3 | Nominally significant missense variants in the 7 AA-AD genes.

Gene	Ch.	BP	rsID	Effect Allele	% AFR	% Cases	% Ctrls	Alt. Allele	*OR*	*P*
ABCA7	19	1,047,336	rs59851484	A	11.88	14.83	10.49	G	1.49	0.0043
	19	1,058,635	rs73505232	T	14.30	16.05	12.18	C	1.38	0.015
	19	1,044,712	rs3764647	G	25.72	26.24	21.60	A	1.29	0.017
	19	1,056,492	rs3752246	G	1.06	3.48	5.72	C	0.59	0.019
	19	1,043,748	rs3752232	G	27.08	27.20	23.20	A	1.24	0.044
	19	1,057,335	rs538930513	A	0.30	0.41	0.00	G	NA	0.049
AKAP9	7	91,732,110	rs149979685	T	0.45	1.13	0.11	C	10.73	0.0046
	7	91,709,085	rs144662445	G	0.53	1.33	0.21	A	6.35	0.0054
	7	91,726,202	rs78351282	A	2.80	3.27	1.59	G	2.10	0.017
	7	91,726,604	rs34956633	G	4.92	5.11	7.54	A	0.66	0.029
	7	91,712,808	rs149946443	A	1.13	0.51	1.48	G	0.34	0.032
	7	91,630,603	rs143894795	C	0.76	0.31	1.06	G	0.29	0.044
SLC10A2	13	103,718,308	rs55971546	T	0.30	0.31	1.06	C	0.29	0.044

Effect allele represents the minor allele; % AFR represents the estimated effect allele frequency in the 1000 Genomes African cohort; % Cases represents the estimated effect allele frequency in AD cases; % Controls represents the estimated effect allele frequency in controls.

study (Logue et al., 2014) (**Table 3**). Our analysis also confirmed the previously reported association for one of the common *ABCA7* missense SNPs noted in N'Songo et al., 2017 (rs3764647, *OR* = 1.29 for minor allele, p = 0.017), but not the rare coding variant (rs3752239, *OR* = 0.39, p = 0.24). Consistent with prior results (Logue et al., 2014), the association with the rare *AKAP9* variants was significant in a mixed model which adjusted for relatedness within the sample with ancestry PCs as covariates (for rs149979685 *OR* = 10.53, p = 0.025 and for rs144662445 *OR* = 6.25, p = 0.016).

Regulatory Variants

We also examined potentially regulatory variants in the AD genes implicated in AAs. Association was tested with variants in regulatory regions for the two primary *AKAP9* isoforms. One variant identified near the TSS of the shorter isoform was not associated with AD (p = 0.66). Significant association was identified with a rare nine bp deletion (rs371245265) located near the TSS for the longer *AKAP9* isoform (*OR* for the deletion = 6.37, p = 0.0053). Prompted by the similarity of allele frequencies between this deletion and the previously identified coding AD risk variants (rs144662445 and rs149979685), we checked the 1000 genomes phase 3 African population data and confirmed high LD between rs371245265 and both rs144662445 (r^2 = 0.86) and rs149979685 (r^2 = 1). Consistent with this information, all 17 discovery sample subjects with the rs371245265 deletion were also carriers of the rs144662445 minor allele, and 14 of these subjects were also carriers of the rs149979685 minor allele. As noted for rs149979685, the association with rs371245265 remained significant in a model adjusting for relatedness and ancestry (*OR* = 6.30, p = 0.016). Nominally significant associations were also observed with three common potentially regulatory SNPs in *ZCWPW1*. The most significant of these three was rs10693652, a 2 bp deletion which was more common in controls than cases (*OR* for the deletion = 0.75, p = 0.0042). The sole *ABCA7* variant and 13 *PTK2B* variants located in TSSs were not associated with AD. Regulatory variants in *COBL* and *SLC10A2* could not be evaluated because the custom capture design for these loci included exons only.

TABLE 4 | Top-ranked association results in previously established AA AD-risk genes.

Gene	Ch.	BP	rsID	Eff. All.	% AFR	% Cases	% Ctrls	Alternate Allele	*OR*	*P*
ABCA7	19	1,050,007	.	–	NA	4.63	2.16	C	2.20	0.0042
	19	1,047,336	rs59851484	A	11.88	14.83	10.49	G	1.49	0.0043
	19	1,049,991	.	–	NA	5.03	2.54	CCTCCCTGT GAGCCCCCC ACCACTT	2.03	0.0065
	19	1,043,260	rs58262414	G	11.72	14.62	10.59	T	1.45	0.0079
	19	1,042,598	rs147599642	A	11.95	14.62	10.70	AAT	1.43	0.0098
AKAP9	7	91,570,040	rs557208555	C	0.53	1.13	0.11	A	10.73	0.0046
	7	91,591,230	rs114789310	A	0.53	1.13	0.11	G	10.73	0.0046
	7	91,663,031	rs183984025	T	0.53	1.13	0.11	C	10.73	0.0046
	7	91,732,110	rs149979685	T	0.45	1.13	0.11	C	10.73	0.0046
	7	91,590,199	rs564709734	G	0.53	1.13	0.11	A	10.70	0.0046
COBL	7	51,085,149	rs150183973	A	1.81	1.94	0.74	T	2.65	0.023
	7	51,098,567	rs142060269	G	NA	44.07	39.07	GTCT	1.23	0.026
	7	51,098,849	rs62448278	A	53.03	48.98	44.17	G	1.21	0.035
	7	51,138,814	rs1295400	T	6.28	7.67	5.30	C	1.49	0.035
MS4A6A	11	59,939,727	rs140130948	G	0.15	0.00	0.64	C	0.00	0.013
	11	59,945,018	rs146080691	A	0.15	0.00	0.32	G	0.00	0.077
	11	59,943,683	rs183204829	T	0.00	0.00	0.32	C	0.00	0.078
	11	59,950,406	rs577683097	A	0.15	0.00	0.32	G	0.00	0.078
	11	59,940,141	rs186332028	C	0.68	0.20	0.74	T	0.27	0.085
PTK2B	8	27,253,935	rs115828696	G	0.68	0.20	1.80	A	0.11	0.00041
	8	27,268,750	rs3757908	T	1.06	3.89	1.91	C	2.08	0.0099
	8	27,276,111	rs891392	C	1.06	3.89	1.91	T	2.08	0.0099
	8	27,272,298	rs144318332	G	4.31	2.05	4.03	C	0.50	0.011
	8	27,280,472	rs77318377	A	4.31	2.05	4.03	G	0.50	0.011
SLC10A2	13	103,718,308	rs55971546	T	0.30	0.31	1.06	C	0.29	0.044
	13	103,718,824	rs16961281	A	13.16	10.84	8.16	G	1.37	0.045
	13	103,719,056	rs7987433	C	23.22	23.21	19.81	T	1.22	0.070
	13	103,697,359	rs199983061	T	0.15	0.20	0.64	C	0.32	0.14
	13	103,697,329	rs41281676	A	4.23	5.11	3.81	G	1.36	0.17
ZCWPW1	7	100,002,772	rs76913697	G	11.57	13.19	18.54	A	0.67	0.0013
	7	100,026,415	rs10693652	TCA	29.43	25.56	31.46	T	0.75	0.0042
	7	100,028,484	rs6962151	C	29.43	25.56	31.46	T	0.75	0.0042
	7	100,025,564	rs67196635	C	29.43	25.61	31.45	T	0.75	0.0047
	7	100,014,313	rs6957928	A	12.93	15.24	19.85	G	0.73	0.0078

Effect allele represents the minor allele; % AFR represents the estimated effect allele frequency in the 1000 Genomes African cohort; % Cases represents the estimated effect allele frequency in AD cases; % Controls represents the estimated effect allele frequency in controls; "–" indicates a deletion; "." indicates a variant without an annotated rsID; NA indicates the variant is not present in 1000 Genomes.

Other Variants

Examination of the full complement of variation in these genes (n = 4,325) including 342 variants in *ABCA7*, 1,445 in *AKAP9*, 167 in *COBL*, 204 in *MS4A6A*, 1,874 in *PTK2B*, 37 variants in *SLC10A2*, and 256 variants in *ZCWPW1* revealed many nominally significant associations (**Table 4**). The most significant association was observed with a rare SNP in *PTK2B* (rs115828696, MAF = 0.0020 in AD cases and 0.18 in controls) which was protective (*OR* for the minor allele A = 0.11, p = 0.00041). A strong protective effect was also identified with a common SNP in *ZCWPW1* (*OR* = 0.67, p = 0.0013). Genotypes were not available for several of the previously implicated AA-specific risk SNPs including *ABCA7* rs115550680 (Reitz et al., 2013a,b) which is located in a repetitive region and was not captured by the design. The *COBL* rs112404845 and *SLC10A2* rs16961023 variants (Mez et al., 2017) are outside of the coding regions and, thus, were not assessed.

Genes Previously Associated With AD in Other Ancestry Groups

Of the 104,341 variants observed in all targeted regions that were tested for association with AD, 29 were annotated as LOF variants. Only the previously noted *MS4A6A* and *ABCA7* variants (rs140130948 and rs567222111) were nominally significant (**Table 5**). The most significant association findings among 1,067 missense variants were obtained with five common highly correlated variants in *PILRB* that showed a protective

TABLE 5 | Top loss of function variants from all sequenced genes (out of 29 LOF variants examined).

Gene	Ch.	BP	rsID	Effect Allele	% AFR	% Cases	% Ctrls	Alt. Allele	*OR*	*P*
MS4A6A	11	59,939,727	rs140130948	G	0.15	0.00	0.64	C	0.00	0.013
ABCA7	19	1,044,707	rs567222111	G	0.83	1.13	0.32	GGGGCACCTGGT	3.57	0.038
CD33	19	51,738,933	rs273621	C	2.27	0.72	1.59	T	0.45	0.072
ACE	17	61,563,661	rs4330	C	41.30	40.11	43.61	A	0.87	0.124
ABCA7	19	1,041,352	rs3752229	G	0.15	0.41	0.95	A	0.43	0.146

Effect allele represents the minor allele; % AFR represents the estimated effect allele frequency in the 1000 Genomes African cohort; % Cases represents the estimated effect allele frequency in AD cases; % Controls represents the estimated effect allele frequency in controls.

TABLE 6 | Top missense variants from all sequenced genes (out of 1,067 missense variants examined).

Gene	Ch.	BP	rsID	Effect Allele	% AFR	% Cases	% Ctrls	Alt. Allele	*OR*	*P*
PILRB	7	99,956,444	rs11761306	G	NA	11.86	17.13	A	0.65	0.0010
PILRB	7	99,956,436	rs11771799	C	7.03	11.89	17.09	T	0.65	0.0012
PILRB	7	99,956,439	rs35986051	C	7.03	11.89	17.09	T	0.65	0.0012
PILRB	7	99,955,866	rs61735533	A	10.14	12.07	17.27	G	0.66	0.0013
ABCA7	19	1,047,336	rs59851484	A	11.88	14.83	10.49	G	1.49	0.0043
AKAP9	7	91,732,110	rs149979685	T	0.45	1.13	0.11	C	10.73	0.0046
AKAP9	7	91,709,085	rs144662445	G	0.53	1.33	0.21	A	6.35	0.0054
KIAA0196	8	126,091,036	rs143719918	T	0.23	1.02	0.11	C	9.74	0.0077
ECHDC3	10	11,797,500	rs35986488	A	3.86	3.78	6.36	G	0.58	0.0100

Effect allele represents the minor allele; % AFR represents the estimated effect allele frequency in the 1000 Genomes African cohort; % Cases represents the estimated effect allele frequency in AD cases; % Controls represents the estimated effect allele frequency in controls; NA indicates the variant is not present in 1000 Genomes.

TABLE 7 | Top potentially regulatory variants from all sequenced genes (out of 223).

Gene	Ch.	BP	rsID	Effect Allele	% AFR	% Cases	% Ctrls	Alt. Allele	*OR*	*P*
ZCWPW1/MEPCE	7	100,026,415	rs10693652	TCA	29.43	25.56	31.46	T	0.75	0.0042
AKAP9	7	91,570,197	rs536714523	T	0.45	1.34	0.21	TGGCGGCGGC	6.37	0.0053
ZCWPW1	7	100,014,846	rs73161762	T	1.13	3.17	5.30	C	0.59	0.020
ZCWPW1/MEPCE	7	100,027,339	rs74460138	G	14.29	16.36	20.44	C	0.76	0.021
SNX6	14	35,099,305	rs562903264	A	0.00	0.51	0.00	G	NA	0.028
KANSL1	17	44,302,765	rs187276691	A	0.15	0.72	0.11	G	6.80	0.038
NSMCE2/KIAA0196	8	126,104,130	rs76575464	A	19.74	18.51	14.97	C	1.29	0.038
CELF1	11	47,574,654	rs575641108	CGCCGCT	0.15	0.50	0.00	C	NA	0.047
BZRAP1/BZRAP1-AS1/MIR142	17	56,406,133	rs374170329	G	0.23	0.41	0.00	C	NA	0.049

Effect allele represents the minor allele; % AFR represents the estimated effect allele frequency in the 1000 Genomes African cohort; % Cases represents the estimated effect allele frequency in AD cases; % Controls represents the estimated effect allele frequency in controls.

effect ($0.0010 < p < 0.0017$; estimated OR for minor alleles varied from 0.65 to 0.67; **Table 6**). Restricting the analysis to potentially regulatory variants, a protective common indel near the TSS of *ZCWPW1* (rs10693652, *OR* for the minor allele = 0.75, $p = 0.0042$) and the rare risk indel near the TSS of *AKAP9* (rs536714523) noted above were the most significant of the 223 variants tested (**Table 7**). Finally, examination of the entire set of 104,341 variants identified in the targeted sequencing experiments yielded significant associations with multiple loci (**Table 8**), most notably a rare protective variant in *F5* (rs2027885, OR for minor allele $A = 0.053$, $p = 6.40 \times 10^{-5}$), a gene that was previously associated with a MRI measure of hippocampal atrophy (Melville et al., 2012), and two common variants in *KIAA0196* ($p < 8.6 \times 10^{-5}$; **Table 8**). Out of the 151 protein-coding genes, nominally significant gene-based associations were found with six genes using the CMC test and with three genes using the VT test (**Table 9**). The most significant of these results was *KANSL1* ($p = 0.013$). None of the seven previously established AD risk genes in AAs were significant ($p > 0.05$).

DISCUSSION

We performed targeted gene sequencing in an AA cohort containing 489 AA AD cases and 472 cognitively normal controls and found evidence of association with several novel variants in genes that were previously implicated with AD risk in AAs including a deletion causing LOF of *ABCA7* (rs567222111). We subsequently genotyped this deletion in an independent cohort

TABLE 8 | Top variants from all genes (out of 104,341 variants examined).

Gene	Ch.	BP	rsID	Eff. All.	% AFR	% Cases	% Ctrls	Alt. Allele	*OR*	*P*
F5	1	169,535,038	rs2027885	A	0.61	0.10	1.91	G	0.053	6.40E-05
KIAA0196	8	126,097,380	rs79300936	A	29.73	30.06	22.14	G	1.51	7.83E-05
KIAA0196	8	126,097,473	rs7832481	G	29.80	29.96	22.08	A	1.51	8.50E-05
KIAA0196	8	126,066,723	rs7817741	A	2.19	6.44	11.12	C	0.55	0.00028
KIAA0196	8	126,073,786	rs2272729	A	2.12	6.44	11.12	G	0.55	0.00028
KIAA0196	8	126,093,882	rs7817303	A	18.31	20.25	14.09	G	1.55	0.00035
PTK2B	8	27,253,935	rs115828696	G	0.68	0.20	1.80	A	0.11	0.00041
PLXNC1	12	94,696,160	rs189295092	–	2.12	5.46	2.29	C	2.46	0.00043
SORL1	11	121,330,087	rs3862606	G	19.82	18.51	25.11	A	0.68	0.00046
PILRB	7	99,965,328	rs11284139	G	10.14	11.96	17.58	GA	0.64	0.00051

Effect allele represents the minor allele; % AFR represents the estimated effect allele frequency in the 1000 Genomes African cohort; % Cases represents the estimated effect allele frequency in AD cases; % Controls represents the estimated effect allele frequency in controls; "–" indicates a single bp deletion.

TABLE 9 | Nominally significant gene-based burden tests of association with AD.

Test	Gene		Start (bp)	End (bp)	Num. Variants Included	*P*
CMC	*KANSL1*	17	44,112,733	44,249,388	124	0.013
	TRIM35	8	27,168,348	27,168,671	12	0.018
	PLEKHM1	17	43,515,240	43,559,893	115	0.029
	MS4A6E	11	60,102,408	60,105,226	9	0.038
	PTK2B	8	27,168,348	27,315,954	163	0.047
	PILRA	7	99,971,735	100,001,863	80	0.053
VT	*CNN2*	19	1,036,201	1,043,455	139	0.017
	PLEKHM1	17	43,515,240	43,559,893	115	0.045
	TRIM35	8	27,168,348	27,168,671	12	0.049

containing 484 AD cases and 484 controls, and the association with this large effect variant (*OR* = 2.42) became more significant in the combined sample. Another notable novel association was identified with a rare 9 bp. deletion (rs371245265) located near the TSS of *AKAP9*. We also confirmed previously reported associations with missense variants in *ABCA7* (rs3764647) and *AKAP9* (rs149979685 and rs144662445). Gene-based tests of aggregated rare variants yielded several associations, most significantly with *KANSL1*, *CNN2*, and *TRIM35*.

The association with the *AKAP9* regulatory region variant rs371245265 calls into question whether the previously identified *AKAP9* missense variants (rs149979685 and rs144662445) are causally related to AD because all of these variants are in high LD. Previous analysis of the background haplotype harboring rs149979685 and rs144662445 and spanning an 800 kb region including five genes showed that no other coding variants could explain the association with these *AKAP9* missense variants (Logue et al., 2014). However, it remains possible that the rs371245265 variant has a regulatory effect on *AKAP9* expression, and this variant alone or in conjunction with the missense variants, could underlie the observed association with AD risk. Because these three rare variants most often co-occur, it is unlikely that the potentially causal effects of these variants will be disentangled by epidemiological studies. Recently, we observed significantly higher phosphorylation and greater post-translational modifications of Tau protein in lymphoblastoid cells from subjects having at least one of the missense variants, a finding that was independent of the disease status of the cell donors (Ikezu et al., 2018). However, since these subjects also have the potentially regulatory variant, experimental studies will be necessary to determine whether this variant does indeed have a regulatory effect and in particular which of the three variants account for the observed effect on Tau phosphorylation.

Our observed novel association with a rare 11 bp loss of function frameshift deletion (rs567222111, Leu396fs) in a gene encoding one of the ATP-binding cassette transporter proteins (*ABCA7*) adds to a growing list of AD-associated LOF mutations in this gene (Farrer, 2015). The most remarkable of these is a 7 bp deletion, causing a frameshift mutation (Glu709fs) that was detected in 11 out of 772 unrelated patients but not in 757 controls from the Flanders region of northern Belgium (Cuyvers et al., 2015). Association of this mutation with AD has also been observed in several other European ancestry populations (Steinberg et al., 2015). Cukier et al reported association of AD and a relatively common 44 bp LOF deletion in *ABCA7* (rs142076058, Ser587fs, *OR* = 2.13) in an AA cohort that is non-overlapping with our study sample (Cukier et al., 2016). This deletion was observed in the current study, but had a smaller effect on AD risk (*OR* = 1.27, $p = 0.16$). Of note, the frameshift mutation identified in our study occurs earlier in the amino acid sequence (position 396) than the Belgian (position 709) or other AA (position 587) frameshift mutations and thus may yield a more seriously impaired protein than these other mutations, but this will have to be confirmed experimentally.

Surprisingly, expanding the analyses from the relatively small set of genes that were implicated in previous studies of AAs to the larger set of AD genes that were established in other populations yielded relatively few significant results, the most significant of which is a rare protective variant (rs2027885) in the gene encoding the blood clotting protein Factor 5 *(F5, OR* = 0.053, $p = 6.40 \times 10^{-5}$). A GWAS of a brain MRI measure of hippocampal atrophy in a MIRAGE Study sample composed primarily of AD and control subjects of European ancestry and a smaller group of AAs (many of which are included in this study) found genome-wide significant association with several common SNPs spanning portions of *F5* and its immediate neighbor, *SELP,* that was supported by evidence in both populations (Melville et al., 2012). Although there is scant genetic evidence linking *F5* to AD, it has been shown that factor V activating protein in Russell's viper venom destabilizes amyloid-β aggregates as revealed from a thioflavin T assay (Bhattacharjee and Bhattacharyya, 2013).

Our findings contrast those of another recent exome sequencing study of AD in a smaller sample of AAs (198 AD cases and 304 controls) which focused exclusively on 20 loci reaching genome-wide significance in a very large GWAS of European ancestry cohorts (N'Songo et al., 2017). The previous study found nominally significant associations with two variants in *ABCA7* (rs3764647 and rs3752239) and in gene-based tests of coding variants in *MS4A6A*, *PTK2B*, and *ZCWPW1*. We observed association with rs3764647 (p = 0.017), but did not replicate the association with rs3752239 or the gene-based associations. On the other hand, gene-based tests of aggregated rare variants in *KANSL1, TRIM35*, *MS4A6E,* and *PILRA* were nominally significant in our study. Differences in findings may be due in part to the use of exome sequencing by N'Songo et al. (2017) versus sequencing of complete gene regions in our study which allowed detection of association with potentially functional variants in regulatory regions and introns that influence transcription and splicing, as well as with structural variants that span non-coding regions.

Our findings should be interpreted cautiously. None of our findings remain significant after correcting for the total number of tests performed in the study. Our sample size was not large enough to detect associations with rare variants exerting modest effects with experiment-wide significance. Also, our primary analyses of individual variants did not account for the correlated structure of our dataset which included many related individuals. Our study highlights the difficulty of obtaining statistically significant results with rare variants, especially those with frequencies less than 1%. It is essential to replicate our findings in independent AA samples, and sufficiently large samples will become available eventually through the efforts of large consortia including the Alzheimer's Disease Genetics Consortium and Alzheimer's Disease Sequencing Project. In addition, experimental studies are needed to establish functionally relevant roles of these genes and variants in AD pathogenesis.

With these concerns in mind, the goal of this study was to identify variants with supporting genetic evidence and predicted functional impact for examination in relevant biological systems. Given the previously identified relationship between loss of function mutations in *ABCA7* and AD (Cuyvers et al., 2015; Farrer, 2015; Steinberg et al., 2015; Cukier et al., 2016) and genetic and biological evidence for a role of rare *AKAP9* variants in AD (Logue et al., 2014; Ikezu et al., 2018), the novel *ABCA7* coding region deletion (rs567222111) and the potentially regulatory *AKAP9* deletion (rs371245265) are the most compelling findings for future studies.

DATA AVAILABILITY

The unprocessed sequence data generated for this study can be found in the National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (https://www.niagads.org/).

AUTHOR CONTRIBUTIONS

MWL, DL, LAF, and KLL contributed to the study design. JF cleaned and processed the sequence data. IS extracted and performed quality control on the DNA specimens used for sequencing and genotyping. MWL and DL performed analyses of the data and prepared the results for presentation. MWL and LAF drafted the manuscript. All authors contributed to the editing and revision of the manuscript.

FUNDING

This study was funded in part by grants from the National Institute on Aging (NIA, R01-AG048927, P30-AG13846, U01-AG032984, UF1-AG046198, and RF1-AG057519). Some biological samples used in this study were obtained from the National Cell Repository for Alzheimer's Disease (NCRAD) at Indiana University funded by NIA grant U24-AG021886.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2018.00592/full#supplementary-material

TABLE S1 | Targeted regions including provisional and confirmed loci associated with AD and related phenotypes.

REFERENCES

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., et al. (2010). A method and server for predicting damaging missense mutations. *Nat. Methods* 7, 248–249. doi: 10.1038/nmeth0410-248

Bhattacharjee, P., and Bhattacharyya, D. (2013). Factor V activator from *Daboia russelli russelli* venom destabilizes beta-amyloid aggregate, the hallmark of Alzheimer disease. *J. Biol. Chem.* 288, 30559–30570. doi: 10.1074/jbc.M113.511410

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:7. doi: 10.1186/s13742-015-0047-8

Chen, H., Wang, C., Conomos, M. P., Stilp, A. M., Li, Z., Sofer, T., et al. (2016). Control for population structure and relatedness for binary traits in genetic

association studies via logistic mixed models. *Am. J. Hum. Genet.* 98, 653–666. doi: 10.1016/j.ajhg.2016.02.012

Chung, J., Wang, X., Maruyama, T., Ma, Y., Zhang, X., Mez, J., et al. (2017). Genome-wide association study of Alzheimer's disease endophenotypes at prediagnosis stages. *Alzheimers Dement* 14, 623–633. doi: 10.1016/j.jalz.2017.11.006

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 6, 80–92. doi: 10.4161/fly.19695

Corder, E. H., Saunders, A. M., Risch, N. J., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., et al. (1994). Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nat. Genet.* 7, 180–184. doi: 10.1038/ng0694-180

Cukier, H. N., Kunkle, B. W., Vardarajan, B. N., Rolati, S., Hamilton-Nelson, K. L., Kohli, M. A., et al. (2016). *ABCA7* frameshift deletion associated with Alzheimer disease in African Americans. *Neurol. Genet.* 2:e79. doi: 10.1212/NXG.0000000000000079

Cuyvers, E., De Roeck, A., Van den Bossche, T., Van Cauwenberghe, C., Bettens, K., Vermeulen, S., et al. (2015). Mutations in *ABCA7* in a Belgian cohort of Alzheimer's disease patients: a targeted resequencing study. *Lancet Neurol.* 14, 814–822. doi: 10.1016/S1474-4422(15)00133-7

Farrer, L. A. (2015). Expanding the genomic roadmap of Alzheimer's disease. *Lancet Neurol.* 14, 783–785. doi: 10.1016/S1474-4422(15)00146-5

Farrer, L. A., Cupples, L. A., Haines, J. L., Hyman, B., Kukull, W. A., Mayeux, R., et al. (1997). Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease meta analysis consortium. *JAMA* 278, 1349–1356.

Farrer, L. A., Sherbatich, T., Keryanov, S. A., Korovaitseva, G. I., Rogaeva, E. A., Petruk, S., et al. (2000). Association between angiotensin-converting enzyme and Alzheimer disease. *Arch. Neurol.* 57, 210–214.

Green, R. C., Cupples, L. A., Go, R., Benke, K. S., Edeki, T., Griffith, P. A., et al. (2002). Risk of dementia among white and African American relatives of patients with Alzheimer disease. *JAMA* 287, 329–336.

Gureje, O., Ogunniyi, A., Baiyewu, O., Price, B., Unverzagt, F. W., Evans, R. M., et al. (2006). APOE epsilon4 is not associated with Alzheimer's disease in elderly Nigerians. *Ann. Neurol.* 59, 182–185. doi: 10.1002/ana.20694

Hendrie, H. C., Hall, K. S., Hui, S., Unverzagt, F. W., Yu, C. E., Lahiri, D. K., et al. (1995). Apolipoprotein E genotypes and Alzheimer's disease in a community study of elderly African Americans. *Ann. Neurol.* 37, 118–120. doi: 10.1002/ana.410370123

Ikezu, T., Chen, C., DeLeo, A. M., Zeldich, E., Fallin, M. D., Kanaan, N. M., et al. (2018). Tau phosphorylation is impacted by rare akap9 mutations associated with Alzheimer Disease in African Americans. *Neuroimmune Pharmacol.* 13, 254–264. doi: 10.1007/s11481-018-9781-x

Jun, G., Asai, H., Zeldich, E., Drapeau, E., Chen, C., Chung, J., et al. (2014). PLXNA4 is associated with Alzheimer disease and modulates tau phosphorylation. *Ann. Neurol.* 76, 379–392. doi: 10.1002/ana.24219

Jun, G., Ibrahim-Verbaas, C. A., Vronskaya, M., Lambert, J. C., Chung, J., Naj, A. C., et al. (2016). A novel Alzheimer disease locus located near the gene encoding tau protein. *Mol. Psychiatry* 21, 108–117. doi: 10.1038/mp.2015.23

Jun, G., Naj, A. C., Beecham, G. W., Wang, L. S., Buros, J., Gallins, P. J., et al. (2010). Meta-analysis confirms CR1, CLU, and PICALM as Alzheimer Disease risk loci and reveals interactions With APOE genotypes. *Arch. Neurol.* 67, 1473–1484. doi: 10.1001/archneurol.2010.201

Jun, G. R., Chung, J., Mez, J., Barber, R., Beecham, G. W., Bennett, D. A., et al. (2017). Transethnic genome-wide scan identifies novel Alzheimer's disease loci. *Alzheimers Dement* 13, 727–738. doi: 10.1016/j.jalz.2016.12.012

Kamboh, M. I., Demirci, F. Y., Wang, X., Minster, R. L., Carrasquillo, M. M., Pankratz, V. S., et al. (2012). Genome-wide association study of Alzheimer's disease. *Transl. Psychiatry* 2:e117. doi: 10.1038/tp.2012.45

Lambert, J. C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., et al. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat. Genet.* 45, 1452–1458. doi: 10.1038/ng.2802

Lee, J. H., Cheng, R., Graff-Radford, N., Foroud, T., and Mayeux, R. (2008). Analyses of the National Institute on aging late-onset Alzheimer's disease family study: implication of additional loci. *Arch. Neurol.* 65, 1518–1526. doi: 10.1001/archneur.65.11.1518

Li, B., and Leal, S. M. (2008). Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. *Am. J. Hum. Genet.* 83, 311–321. doi: 10.1016/j.ajhg.2008.06.024

Logue, M. W., Schu, M., Vardarajan, B. N., Buros, J., Green, R. C., Go, R. C., et al. (2011). A comprehensive genetic association study of Alzheimer disease in African Americans. *Arch. Neurol.* 68, 1569–1579. doi: 10.1001/archneurol.2011.646

Logue, M. W., Schu, M., Vardarajan, B. N., Farrell, J., Bennett, D. A., Buxbaum, J. D., et al. (2014). Two rare AKAP9 variants are associated with Alzheimer's disease in African Americans. *Alzheimers Dement* 10, 609.e11–618.e11. doi: 10.1016/j.jalz.2014.06.010

MacArthur, D. G., Balasubramanian, S., Frankish, A., Huang, N., Morris, J., Walter, K., et al. (2012). A systematic survey of loss-of-function variants in human protein-coding genes. *Science* 335, 823–828. doi: 10.1126/science.1215040

McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R. Jr., Kawas, C. H., et al. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 7, 263–269. doi: 10.1016/j.jalz.2011.03.005

Meier, I. B., Manly, J. J., Provenzano, F. A., Louie, K. S., Wasserman, B. T., Griffith, E. Y., et al. (2012). White matter predictors of cognitive functioning in older adults. *J. Int. Neuropsychol. Soc.* 18, 414–427. doi: 10.1017/S1355617712000227

Melville, S. A., Buros, J., Parrado, A. R., Vardarajan, B., Logue, M. W., Shen, L., et al. (2012). Multiple loci influencing hippocampal degeneration identified by genome scan. *Ann. Neurol.* 72, 65–75. doi: 10.1002/ana.23644

Meng, Y., Baldwin, C. T., Bowirrat, A., Waraska, K., Inzelberg, R., Friedland, R. P., et al. (2006). Association of polymorphisms in the angiotensin-converting enzyme gene and Alzheimer disease in an Israeli-Arab community. *Am. J. Hum. Genet.* 78, 871–877. doi: 10.1086/503687

Mez, J., Chung, J., Jun, G., Kriegel, J., Bourlas, A. P., Sherva, R., et al. (2017). Two novel loci, COBL and SLC10A2, for Alzheimer's disease in African Americans. *Alzheimers Dement* 13, 119–129. doi: 10.1016/j.jalz.2016.09.002

N'Songo, A., Carrasquillo, M. M., Wang, X., Burgess, J. D., Nguyen, T., Asmann, Y. W., et al. (2017). African American exome sequencing identifies potential risk variants at Alzheimer disease loci. *Neurol. Genet.* 3:e141. doi: 10.1212/NXG.0000000000000141

Price, A. L., Kryukov, G. V., de Bakker, P. I., Purcell, S. M., Staples, J., Wei, L. J., et al. (2010). Pooled association tests for rare variants in exon-resequencing studies. *Am. J. Hum. Genet.* 86, 832–838. doi: 10.1016/j.ajhg.2010.04.005

Reitz, C., Jun, G., Naj, A., Rajbhandary, R., Vardarajan, B. N., Wang, L. S., et al. (2013a). Variants in the ATP-binding cassette transporter (*ABCA7*), apolipoprotein E 4,and the risk of late-onset Alzheimer disease in African Americans. *JAMA* 309, 1483–1492. doi: 10.1001/jama.2013.2973

Reitz, C., Tosto, G., Vardarajan, B., Rogaeva, E., Ghani, M., Rogers, R. S., et al. (2013b). Independent and epistatic effects of variants in VPS10-d receptors on Alzheimer disease risk and processing of the amyloid precursor protein (APP). *Transl. Psychiatry* 3:e256. doi: 10.1038/tp.2013.13

Rogaeva, E., Meng, Y., Lee, J. H., Gu, Y.-J., Zou, F., Kawarai, T., et al. (2007). The sortilin-related receptor SORL1 is functionally and genetically associated with Alzheimer's disease. *Nat. Genet.* 39, 168–177.

Sahota, A., Yang, M., Gao, S., Hui, S. L., Baiyewu, O., Gureje, O., et al. (1997). Apolipoprotein E-associated risk for Alzheimer's disease in the African-American population is genotype dependent. *Ann. Neurol.* 42, 659–661. doi: 10.1002/ana.410420418

Saunders, A. M., Strittmatter, W. J., Schmechel, D., George-Hyslop, P. H., Pericak-Vance, M. A., Joo, S. H., et al. (1993). Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 43, 1467–1472.

Steinberg, S., Stefansson, H., Jonsson, T., Johannsdottir, H., Ingason, A., Helgason, H., et al. (2015). Loss-of-function variants in *ABCA7* confer risk of Alzheimer's disease. *Nat. Genet.* 47, 445–447. doi: 10.1038/ng.3246

Vardarajan, B. N., Bruesegem, S. Y., Harbour, M. E., Inzelberg, R., Friedland, R., St George-Hyslop, P., et al. (2012). Identification of Alzheimer disease-associated variants in genes that regulate retromer function. *Neurobiol. Aging* 33, 2231.e15–2231.e30. doi: 10.1016/j.neurobiolaging.2012.04.020

Wetzel-Smith, M. K., Hunkapiller, J., Bhangale, T. R., Srinivasan, K., Maloney, J. A., Atwal, J. K., et al. (2014). A rare mutation in UNC5C predisposes to late-onset Alzheimer's disease and increases neuronal cell death. *Nat. Med.* 20, 1452–1457. doi: 10.1038/nm.3736

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

METHODS
published: 28 August 2018
doi: 10.3389/fnins.2018.00589

New Antibody-Free Mass Spectrometry-Based Quantification Reveals That C9ORF72 Long Protein Isoform Is Reduced in the Frontal Cortex of Hexanucleotide-Repeat Expansion Carriers

*Arthur Viodé[1], Clémence Fournier[2,3], Agnès Camuzat[2,4], François Fenaille[1], NeuroCEB Brain Bank, Morwena Latouche[2,4], Fanny Elahi[5], Isabelle Le Ber[2,3,6], Christophe Junot[1], Foudil Lamari[7,8], Vincent Anquetil[2,3] and François Becher[1]**

[1] Service de Pharmacologie et Immunoanalyse, Laboratoire d'Etude du Métabolisme des Médicaments, Commissariat à l'Énergie Atomique et aux Énergies Alternatives, Institut National de la Recherche Agronomique, Université Paris Saclay, Gif-sur-Yvette, France, [2] Inserm U 1127, CNRS UMR 7225, Institut du Cerveau et de la Moelle Épinière, ICM, Sorbonne Université, Paris, France, [3] Assistance Publique – Hôpitaux de Paris, Hôpital Pitié-Salpêtrière, Paris, France, [4] Ecole Pratique des Hautes Etudes, PSL Research University, Paris, France, [5] Department of Neurology, Memory and Aging Center, University of California, San Francisco, San Francisco, CA, United States, [6] National Reference Center for Rare or Early Dementias, Institute of Memory and Alzheimer's Disease (IM2A), Department of Neurology, AP-HP – Hôpital Pitié-Salpêtrière, Paris, France, [7] Assistance Publique – Hôpitaux de Paris, Service de Biochimie Métabolique, Hôpitaux Universitaires Pitié Salpêtrière – Charles Foix, Paris, France, [8] GRC 13 Neurométabolisme – UPMC, Sorbonne Université, Paris, France

OPEN ACCESS

Edited by:
Leo P. Sugrue,
University of California,
San Francisco, United States

Reviewed by:
Chiara F. Valori,
Helmholtz-Gemeinschaft Deutscher Forschungszentren (HZ), Germany
Dapeng Chen,
University of Maryland, College Park, United States

***Correspondence:**
François Becher
francois.becher@cea.fr

Specialty section:
This article was submitted to Neurodegeneration, a section of the journal Frontiers in Neuroscience

Received: 15 March 2018
Accepted: 06 August 2018
Published: 28 August 2018

Citation:
Viodé A, Fournier C, Camuzat A, Fenaille F, NeuroCEB Brain Bank, Latouche M, Elahi F, Le Ber I, Junot C, Lamari F, Anquetil V and Becher F (2018) New Antibody-Free Mass Spectrometry-Based Quantification Reveals That C9ORF72 Long Protein Isoform Is Reduced in the Frontal Cortex of Hexanucleotide-Repeat Expansion Carriers. Front. Neurosci. 12:589. doi: 10.3389/fnins.2018.00589

Frontotemporal dementia (FTD) is a fatal neurodegenerative disease characterized by behavioral and language disorders. The main genetic cause of FTD is an intronic hexanucleotide repeat expansion (G_4C_2)n in the *C9ORF72* gene. A loss of function of the C9ORF72 protein associated with the allele-specific reduction of *C9ORF72* expression is postulated to contribute to the disease pathogenesis. To better understand the contribution of the loss of function to the disease mechanism, we need to determine precisely the level of reduction in C9ORF72 long and short isoforms in brain tissue from patients with *C9ORF72* mutations. In this study, we developed a sensitive and robust mass spectrometry (MS) method for quantifying C9ORF72 isoform levels in human brain tissue without requiring antibody or affinity reagent. An optimized workflow based on surfactant-aided protein extraction and pellet digestion was established for optimal recovery of the two isoforms in brain samples. Signature peptides, common or specific to the isoforms, were targeted in brain extracts by multiplex MS through the parallel reaction monitoring mode on a Quadrupole–Orbitrap high resolution mass spectrometer. The assay was successfully validated and subsequently applied to frontal cortex brain samples from a cohort of FTD patients with *C9ORF72* mutations and neurologically normal controls without mutations. We showed that the C9ORF72 short isoform in the frontal cortices is below detection threshold in all tested individuals and the C9ORF72 long isoform is significantly decreased in *C9ORF72* mutation carriers.

Keywords: frontotemporal dementia (FTD), frontotemporal lobar degeneration (FTLD), amyotrophic lateral sclerosis (ALS), C9ORF72, TDP-43, TDP43, mass spectrometry (MS), GRN

INTRODUCTION

Frontotemporal dementia (FTD) is the second most prevalent neurodegenerative disease before the age of 65, after Alzheimer disease. FTD are caused by frontal and temporal lobar degeneration leading to behavioral, socioemotional, language disorders, and progressive loss of autonomy and death approximately 10 years after disease onset (Rascovsky et al., 2011). Amyotrophic lateral sclerosis (ALS) caused by motor neuron degeneration is associated with FTD in 15% of patients or families. Familial forms of FTD, accounting for about 20–50% of cases (Rosso, 2003) are mainly caused by mutations in three major genes: granulin (*GRN*), microtubule-associated protein tau (*MAPT*) and C9 open reading frame 72 (*C9ORF72*).

C9ORF72, the most frequent genetic etiology, represents 25% of familial FTD cases (DeJesus-Hernandez et al., 2011; Renton et al., 2011), 80% of FTD-ALS familial cases and also explains 40% of pure familial ALS (without FTD) (Majounie et al., 2012). An expansion of hundreds to thousands hexanucleotide (G_4C_2) repeats is present in the first intron of the gene in patients, while healthy individuals carry less than 24 (G_4C_2) repeats. While the exact mechanisms of disease remain unknown, two main modes of toxicity are proposed. The expansion may be deleterious through formation of nuclear RNA foci by RNA containing the hexanucleotide expansion (G_4C_2) with sequestration of RNA-binding proteins (Lagier-Tourenne et al., 2013), and/or translation of polydipeptides proteins that aggregate in the brain (Mori et al., 2013). A loss of function of the C9ORF72 protein associated with a decrease in *C9ORF72* allele-specific expression was also hypothesized to contribute to the pathogenesis based on reduction of *C9ORF72* mRNA transcript levels in patients (Ciura et al., 2013). The three alternatively spliced *C9ORF72* transcripts encode two C9ORF72 protein isoforms, a 222 amino acids (AA) protein isoform called C9-short (C9-S) and a 481AA protein isoform called C9-long (C9-L). The two protein isoforms have been shown by immunofluorescence to have distinct cellular localization with the C9-S lozalizing to the nuclear membrane and C9-L to the cytoplasm (Xiao et al., 2016). This observation suggests that the two isoforms have a different function, while the precise function of the protein has not been clearly determined. A potential role in endosomal transport and autophagy was reported through interaction with Rab-GTPases (Zhang et al., 2012; Levine et al., 2013).

A better understanding of the contributions of the loss of function to the disease mechanism requires a precise quantification of reduction in levels of C9ORF72 isoforms. C9ORF72 protein has been so far only quantified in human tissues by Western blot (Waite et al., 2014). Several studies pointed out the poor affinity and selectivity of commercially available antibodies (Waite et al., 2014; Davidson et al., 2018), requiring laborious in-house generation of antibodies. Targeted mass spectrometry (MS) is a powerful alternative to Western blot and enzyme-linked immunosorbent assay (ELISA) for quantification of proteins. It provides accurate quantification, high level of specificity, avoiding interference due to cross-reactivity of antibodies, and the ability to discriminate between isoforms (Chen D. et al., 2015; Chen Y.-T. et al., 2015; Jedrychowski et al., 2015; Lesur et al., 2015). Nevertheless, MS-based detection of low-abundant proteins in complex fluids or tissues remains challenging without efficient sample preparation protocols. The gold standard relies on the combination with immunoprecipitation to selectively enrich the analyte of interest prior to MS (Chen Y.-T. et al., 2015), but is applicable only when antibodies with sufficient specificity and affinity for the target protein are available.

We have developed a sensitive and robust antibody-free MS assay for quantification of C9ORF72 isoforms in brain samples. The protocol consists in an optimized tissue lysis protocol followed by pellet digestion of extracted brain proteins and specific monitoring of common and isoform specific peptides by targeted high-resolution MS in the parallel reaction monitoring mode (PRM). Reproducibility and linearity were demonstrated, as well as equivalent isoform recovery from brain tissue samples and stability during sample preparation. This new assay allowed for the first time the quantification of the C9ORF72 long isoform in *post mortem* frontal cortex brain samples from a cohort of FTD patients harboring a *C9ORF72* mutation and highlighted a significant decrease in concentrations in mutation carriers. The short isoform was found to be below the sensitivity threshold of the method.

MATERIALS AND METHODS

Patients Information, Tissue Collection, and Consent

Frozen tissue from frontal cortex (Brodmann area 9/10) of 21 FTD (with or without secondarily developed ALS) patients carrying *C9ORF72* expansion, of 10 patients with non-genetic FTD (with or without secondarily developed ALS) pathologically characterized by TDP-43-positive neuronal inclusions, and of 12 neurologically healthy controls were studied. The brain samples were collected through a brain donation program dedicated to neurodegenerative dementias coordinated by the NeuroCEB Brain Bank Network. The informed consent for post-mortem examination and research studies was signed by the legal representative of each patient in patient's name, as allowed by the French law and approved by the local ethics committee and the brain bank has been officially authorized to provide samples to scientists (agreement AC-2013-1887). All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional research committees and with the 1964 Helsinki declaration. The brain banks fulfill criteria from the French Law on biological resources including informed consent, ethics review committee and data protection (article L1243-4 du Code de la Santé publique, August 2007). The Neuro-CEB brain bank (BioResource Research Impact Factor number BB-0033-00011) has been declared to the Ministry of Research and Higher Education, as required by French law.

Chemical and Materials

C9ORF72 short isoform was purchased from Proteintech (cat# ag21080) (Proteintech Group, Chicago, IL, United States)

and C9ORF72 long isoform from Abnova (cat#00203228-P01) (Abnova Le Perray En Yvelines, France). Trypsin from bovine pancreas TPCK Treated (reference T1426), ammonium hydroxide, ammonium bicarbonate were purchased from Sigma-Aldrich (Saint Quentin Fallavier, France). RapiGest SF Surfactant and SPE Oasis Max 1CC/30 mg were purchased from Waters Corporation (Milford, MA United States). Labeled peptides for quantification were synthesized in Absolute QUAntitation (AQUA) ultimate quality by Thermo Fisher Scientific (Paisley, United Kingdom). Water (ChromaSolve LC-MS), acetonitrile (HPLC-grade), and formic acid were obtained from Honeywell/Riedel-de Haen (Seelze, Germany) and VWR chemicals (Fontenay sous Bois, France), respectively. All other chemicals were purchased from Sigma-Aldrich (Saint Quentin Fallavier, France) or VWR Chemicals (Fontenay sous Bois, France). Pierce BCA protein Assay kit was purchased from Pierce (Rockford, IL, United States). For all reactions, LoBind Eppendorf tubes (Dutscher, Brumath, France) were used.

Sample Preparation

Brain Protein Extraction Protocol

Lysis buffer containing, trizma-base 20 mM; NaCl 150 mM; cOmplete Protease Inhibitor Cocktail 1X and 1% triton, was added to single pieces of whole brain tissue (∼100 mg) at a ratio of 5 μL per 1 mg of tissue. Brain samples were homogenized by beads beating using a precellys soft tissue CK14 2 mL (3*30S at 6,500 rpm). The lysate was then centrifuged at 4,000 rpm for 15 min at 4°C. Fifty microliters from the supernatant was used for analysis. A 5 μL aliquot was used for total protein concentration, determined by the Pierce BCA Protein Assay kit with a sample to working reagent ratio 1:20. Two percent SDS were added to eliminate interference from lipids.

Pellet Digestion

The lysate was precipitated by adding 150 μL of methanol (ratio 3:1), followed by vortex-mixing and briefly centrifuged 5 s. The supernatant was discarded. Twenty microliters of rapigest 0.05% in ammonium bicarbonate 50 mM were added to the pellet. Aqua peptides were added at this step, at 10 ng/mL final concentration. After mixing for approximately 15 min, reduction was performed with 10 μL DTT (20 mM) and incubation at 60°C for 30 min. Alkylation was performed with 10 μL iodoacetamide (45 mM) and 45 min incubation at room temperature. Proteins were digested overnight at 37°C with 40 μg of trypsin.

Solid Phase Extraction (SPE)

Tryptic digests were diluted by addition of 300 μL of 5% ammonium hydroxide before SPE extraction on oasis MAX 1 cc/30 mg, previously conditioned with 1 mL of Methanol and equilibrated with 1 mL of water. Samples were loaded and washed with 500 μL of 5% NH40H and 2*250 μL of methanol. Peptides were eluted with 3*250 μL of methanol containing 10% formic acid. Extracts were then evaporated to dryness with a Turbovap instrument (Biotage, Uppsala, Sweden) (5–15 psi, 40°C for 1 h). The dry residue was re-dissolved in 95% water 5% acetonitrile 0.1% formic acid and centrifuged at 4°C for 10 min at 15,000 × *g*, prior injection into the LC system.

LC-MS/MS Analysis

LC-MS/MS was performed on a Dionex Ultimate 3,000 chromatography system coupled to a Q-exactive Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). Ten microliters of sample was loaded onto the column. Chromatographic separation was performed on an Aeris peptide XBC18 reverse phase column (150 mm × 2.1 mm; 1.7 μm; 100 Å; phenomenex, Le Pecq, France) and achieved in a 30 min gradient at a flow rate of 500 μL/min. A gradient of mobile phase consisting of LC-MS-grade water with 0.1% formic acid (phase A) and acetonitrile with 0.1% formic acid (phase B) was delivered. After an isocratic step of 0.5 min at 5% B, the gradient was ramped to 25% over the next 19.5 min then to 50% over the next 4min. Then acetonitrile was increased to 95% for the next 2 min. Column re-equilibration at 5% B was realized for 4 min.

Instrument parameters of the electrospray ionization source were set as follows: sheath gas flow rate at 70 a.u., spray voltage at 4 kV, capillary temperature at 320°C. The Q-exactive instrument was operated in positive ion mode under time-scheduled sequential PRM acquisition. Endogenous peptides precursor ions and AQUA peptides were selected in the quadrupole with an isolation mass window of 1.5 m/z. Precursors were fragmented in the HCD cell using nitrogen as collision gas and the optimized normalized collision energy (**Supplementary Table S2**). All fragment ions were transferred to the Orbitrap. Resolution was set to 70,000 at m/z 200 (full width at half-maximum), automatic gain control to 1e6, and maximum injection time to 240 ms. Xcalibur 2.2 software (Thermo Fisher Scientific, Bremen, Germany) was used for instrument control and processing of the data files.

PRM and Quantification

A time-scheduled sequential PRM method was established targeting the following C9ORF72 peptides TEIALSGK, ILLEGTER, DSTGSFVLPFR, and SHSVPEEIADIADTVLNDDDI GDSCHEGFLLK (**Supplementary Table S2**). To increase the signal to noise ratio and assay sensitivity, the signal of up to 6 major and non-interfered fragment ions identified with high resolution (5 ppm) from a common peptide precursor were summed up to provide one extracted ion chromatogram (XIC) for each peptide. Isotope-labeled synthetic peptides with labeled amino acids $^{13}C_6$,$^{15}N_2$-labeled lysine and $^{13}C_6$,$^{15}N_4$-labeled arginine were used for signal normalization and quantification of C9ORF72 peptides. Raw MS data were exported to Skyline 3.7 (MacLean et al., 2010) for verification of the transitions ratio of unlabeled and labeled peptides. Xcalibur 2.2 software (Thermo Fisher Scientific, Bremen, Germany) was used for quantitative data analysis.

External calibration curve was made by digesting the C9ORF72 long isoform recombinant protein in a surrogate matrix (mice brain) and SIL peptides were spiked before reduction and alkylation of the pellet. Linear regression with 1/x weighting was applied to generate a standard curve.

HEK293 Cells Transfection

Plasmids expressing EGFPN-tagged C9ORF72 long or short isoforms were assembled as follow. cDNAs coding for NP_060795 (long) and NP_659442 (short) were ordered from DNA2.0/ATUM in the pCS2 vector. Fifty nanograms of each plasmid was PCR amplified with AccuPrime Pfx Supermix (Invitrogen) according to the manufacturer's protocol. Forward primer was CACCTCGACTCTTTGCCCACC and reverse primers were, respectively, for the long and short isoforms CTAAAAAGTCATTAGAACATCTCGTTCTTGCACAC and CTACTTGAGAAGAAAGCCTTCATGACAGC. Purified and sequenced PCR products were cloned into pENTR/D Gateway TOPO according to manufacter's protocol (Thermo Fisher Scientific). Purified and sequenced entry long and short C9ORF72 clones were recombined with pgLAP1 destination vector (Addgene Plasmid #19702) with LR Clonase Enzyme Mix according to the protocol given by the manufacturer (Thermo Fisher Scientific). After sequencing, the two plasmids, EGFP long C9ORF72 and EGFP short C9ORF72, were transfected with lipofectamine 2000 (Thermo Fisher Scientific) according to manufacturer's protocol. Briefly, T-75 flasks of HEK293 cells at 70% confluency were transfected with 15 μg EGFP long C9ORF72 and 15 μg EGFP plasmids, or 15 μg EGFP short C9ORF72 and 15 μg EGFP plasmids or 15 μg EGFP long C9ORF72 and 15 μg EGFP short C9ORF72 plasmids, or 30 μg EGFP control plasmid. Cells were collected 48 h after transfection. PBS washed pellets were stored at minus 80°C until protein extraction.

Statistical Analyses

All statistical analysis was performed using Graphpad Prism software (version 5.01). Data were compared with a Mann–Whitney test and medians were considered significantly different if $p < 0.05$. Data were represented with median and interquartile range.

RESULTS

Assay Design

First step of the assay (**Figure 1**) consists in an optimized brain sample homogenization and extraction of C9ORF72 isoforms using bead-beating tubes with the presence of a nonionic surfactant Triton X-100 to further disrupt lipidic cell membranes. Then, a pellet digestion protocol was adapted to protein digestion in brain extracts. C9ORF72 isoforms are denaturated and equivalently precipitated with methanol. After removal of the supernatant containing the surfactant, resuspension and addition of AQUA peptides, i.e., stable isotope peptide, the pellet was digested with trypsin. Signature peptides were then extracted using mixed-mode anion exchange cartridges. Finally, after evaporation and resuspension in 5% acetonitrile with 0.1% formic acid, samples were injected into the LC-MS/MS where three peptides, common to both isoforms or unique to the long isoform, were quantified with specificity in the PRM mode. A specific peptide from the short isoform was also monitored for detection purposes only.

Analytical Procedure

Proteotypic Peptides Selection and LC-MS/MS Detection

Two isoforms of the C9ORF72 protein are reported (DeJesus-Hernandez et al., 2011) (Uniprot sequence entries: Q96LT7-1 and Q96LT7-2). Amino-acid sequences are illustrated in **Figure 2**. Short isoform amino acid sequence is shared with the long one with the exception of its last residue, i.e., residue 222 (N→K). For a comprehensive assay of C9ORF72, common and isoform-specific peptides have to be identified. To this end, data-dependent analysis (DDA) experiments on human frontal cortices were performed (described in Supporting Information). No signal was detected for the C9ORF72 isoforms, similarly, to another recent study (Umoh et al., 2018), illustrating their low abundance in brain and the need to develop a more sensitive targeted method. A tryptic digest of recombinant C9ORF72 long isoform was then used for selection of best responding peptides. Three peptides were selected based on signal intensity after digestion, size (i.e., between 6 and 20 amino acids), and lack of cysteine, methionine and glutamine residues. Specificity of the selected peptides was assessed using Basic Local Alignment Search Tool (BLAST) against the UniprotKB/Swissprot human database. Among the three peptides selected for quantification (**Figure 2**), two are common to both the short and long isoforms and one is unique to the long isoform. The assay based on peptide 1 (TEIALSGK) and 2 (IILEGTER) quantify total

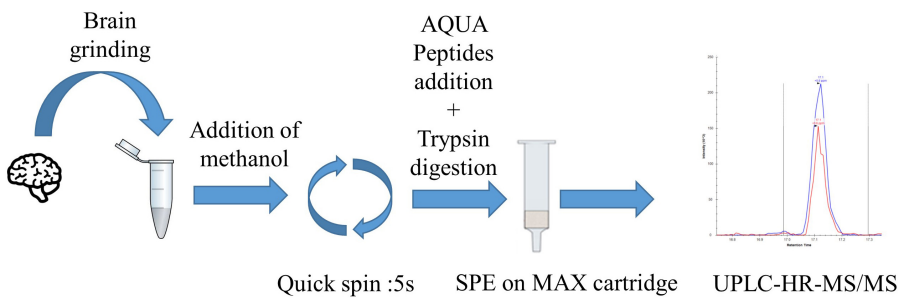

FIGURE 1 | C9ORF72 quantification workflow by pellet digestion and LC-MS/MS in brain samples. First step consists of tissue lysis followed by brain proteins precipitation. Then, the pellet is digested by trypsin. Peptides are extracted and concentrated by SPE before LC-MS/MS injection.

FIGURE 2 | C9ORF72 long (top) and short (bottom) isoform (respectively, Q96LT7-1 and Q96LT7-2). Proteotypic peptids selected are underlined. A red box displays the single amino acid differentiating the common sequence.

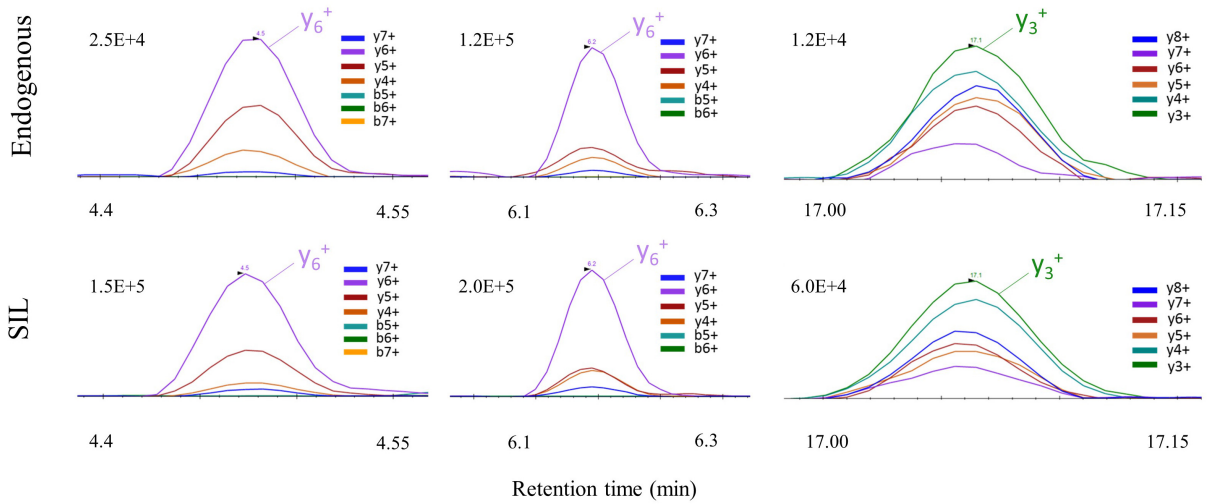

FIGURE 3 | PRM signal of C9ORF72 proteotypic peptides with their corresponding co-eluting SIL peptide in human brain samples.

C9ORF72 whereas the assay based on peptide 3 (DSTGSFVLPFR) differentiates the long isoform. Internal standards AQUA peptides of the three selected peptides for quantification were synthesized by incorporating stable isotopes at the C-terminal amino acid residues ($^{13}C_6$,$^{15}N_2$-labeled lysine and $^{13}C_6$,$^{15}N_4$-labeled arginine). AQUA peptides were prepared with high purity (>95%) and well-defined concentrations. The unique 30 amino acids long peptide specific for the short isoform (peptide 4), containing residue 222 (**Figure 2**), did not meet the stringent selection criteria for inclusion in the quantitative method, and was selected for detection only. A stable isotope-labeled (SIL) analog of peptide 4 was nonetheless synthetized for unambiguous identification of C9-S in the brain samples. In the absence of a suitable quantitative peptide for C9-S, quantification is obtained by difference as previously reported for progastrin-releasing peptide (ProGRP) isoforms (Torsetnes et al., 2013).

Simplification of sample preparation oriented our choice toward Parallel Reaction Monitoring (PRM) over Selected Reaction Monitoring (SRM) for the increased specificity regarding fragment ion detection in complex matrixes provided by the high-resolution Orbitrap mass analyzer. A PRM method was established targeting the three selected peptides for quantification, and the unique C9-S peptide for detection only. Ultra-high performance liquid chromatography was performed on a C18 column for peptide separation with a total runtime of 30 min. Shorter gradients resulted in interferences at the retention time of peptide DSTGSFVLPFR (SIL version) and a decrease of signal intensity up to a factor of 2 in brain samples (**Supplementary Figure S1** and **Supplementary Table S1**). Data treatment increased sensitivity by summing the signals of up to six major and non-interfered fragment ions to provide one XIC for each targeted peptide (Dupré et al., 2015; **Figure 3** and **Supplementary Table S2**). Each endogenous peptide and their corresponding isotope-labeled form (AQUA peptides) must strictly co-elute with similar transition ratio across the different samples in comparison with a standard. If not, the transition was excluded.

Ability of the targeted method to detect the isoforms was demonstrated in HEK293 cells transfected with plasmids expressing either C9-S or C9-L. All common (peptides 1 and 2) and the two unique isoform-specific peptides of C9-L and C9-S (peptides 3 and 4) were detected in corresponding HEK293 cells (**Supplementary Figure S3**). Mock-transfected HEK293 cells were also analyzed and displayed a lower level of C9ORF72.

Sample Preparation of Human Brain Samples Prior to LC-PRM

The main steps of sample preparation consist of tissue homogenization for optimal protein extraction followed by a pellet digestion and peptide enrichment with SPE for lowering ion suppression/matrix effects and increasing sensitivity. Protein extraction protocols with tissue homogenization by mechanical shear recently published for C9ORF72 (Xiao et al., 2015), containing either low/high-salt content (i.e., 150 mM NaCl or 750 mM NaCl), 8 M urea or Triton X-100, were individually evaluated on tissues samples from frontal cortex (**Figure 4**). These protocols were adapted to single use

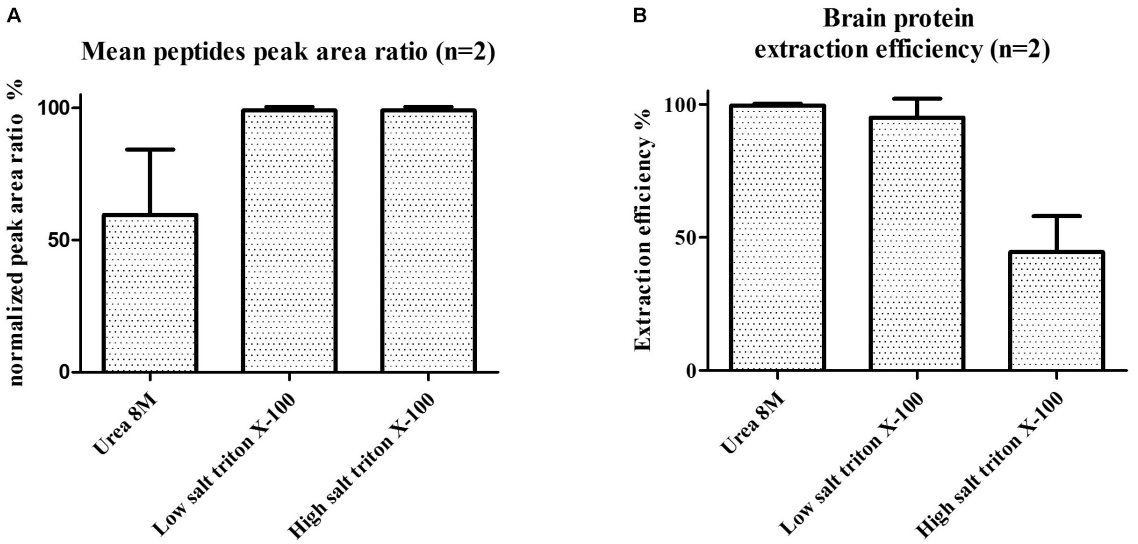

FIGURE 4 | (A) Mean C9ORF72 peptides signal in human brain samples and **(B)** brain protein extraction efficiency (measured by BCA assay) depending on tissue lysis protocols. Protocols evaluated contained either 8 M urea, low salt (150 mM NaCl) triton X-100 or high salt (750 mM NaCl) triton X-100.

bead-beating Precellys tubes to parallelize protein extraction and avoid contamination between samples. Protocols were evaluated based on total protein extraction and signal from C9ORF72 peptides. The low salt protocol without detergent resulted in low protein extraction yield and no signal for C9ORF72 peptides, although Western blot signal was previously reported for the short isoform (Xiao et al., 2015). Signals were observed for the three quantitative peptides selected for quantification (peptides 1–3) when applying the Triton containing mixtures, the latter showing higher intensity than that containing urea (**Figure 4A**). We finally selected the low-salt Triton protocol because it allowed the extraction of more total proteins than the high-salt Triton protocol (**Figure 4B**). However, no signal was detected for the C9-S unique peptide (peptide 4) with any of the evaluated protocols.

A recently published convenient pellet digestion protocol for monoclonal antibody quantification in human plasma (Becher et al., 2017) was evaluated for the detection and quantification of C9ORF72 isoforms in brain proteins lysates. Recombinant protein of C9-S like and C9-L spiked into mice brain lysate, as a substitute to human brain, before and after methanolic precipitation demonstrated high and equivalent recovery for both recombinant protein isoforms with equivalent precipitation yield above 70% and coefficients of variation below 10% (n = 2) (**Table 1**). Brain lysate is a highly complex matrix resulting in deleterious matrix effects. To further reduce its complexity, SPE clean-up was evaluated in human brain samples based on AQUA peptides signal after SPE (**Supplementary Figure S2**). Among Oasis Hydrophilic-Lipophilic Balance (HLB) operated at high pH, Mixed-mode Cation-eXchange (MCX), and Mixed-mode Anion-eXchange (MAX) cartridges, Oasis MAX has proven to give the higher peptide signal intensity (data not shown), in line with the acidic isoelectric point of the peptides ranging from 3.9 to 5.8. Signal enhancement resulting from SPE was about two folds. SPE yield was then determined in the final conditions in mice brain lysate. Yield measured by aqua peptides spiked before and after the SPE extraction were between 45 and 70% (**Supplementary Table S3**), which is in line with previous works (Gong et al., 2015). For method robustness, aqua peptides were spiked early in the protocol, before trypsin digestion of the pellet, and therefore extracted by SPE, similarly, to the C9ORF72 peptides.

TABLE 1 | Precipitation yield and variation of C9-S and C9-L isoforms in mice brain extracts.

	Short isoform		Long isoform	
	Mean	CV	Mean	CV
Precipitation yield (n = 2)	88%	8%	74%	9%

Method Validation

The assay was then evaluated for C9ORF72 quantification in human brain tissues, including the main items of linearity, sensitivity, inter-, and intra-day assay precision, stability during sample processing, i.e., over 90 min at room temperature and matrix effect. Linearity and sensitivity were determined with standard curves prepared in mice brain as a surrogate to human brain samples. Indeed, the three proteotypics peptides selected for quantification are not present in mice due to single point mutations, despite 85% sequence identity with human (**Supplementary Figure S4**). Other items were tested with quality control samples prepared in human brain extract, spiked with recombinant C9-L. Recombinant C9-L was selected for

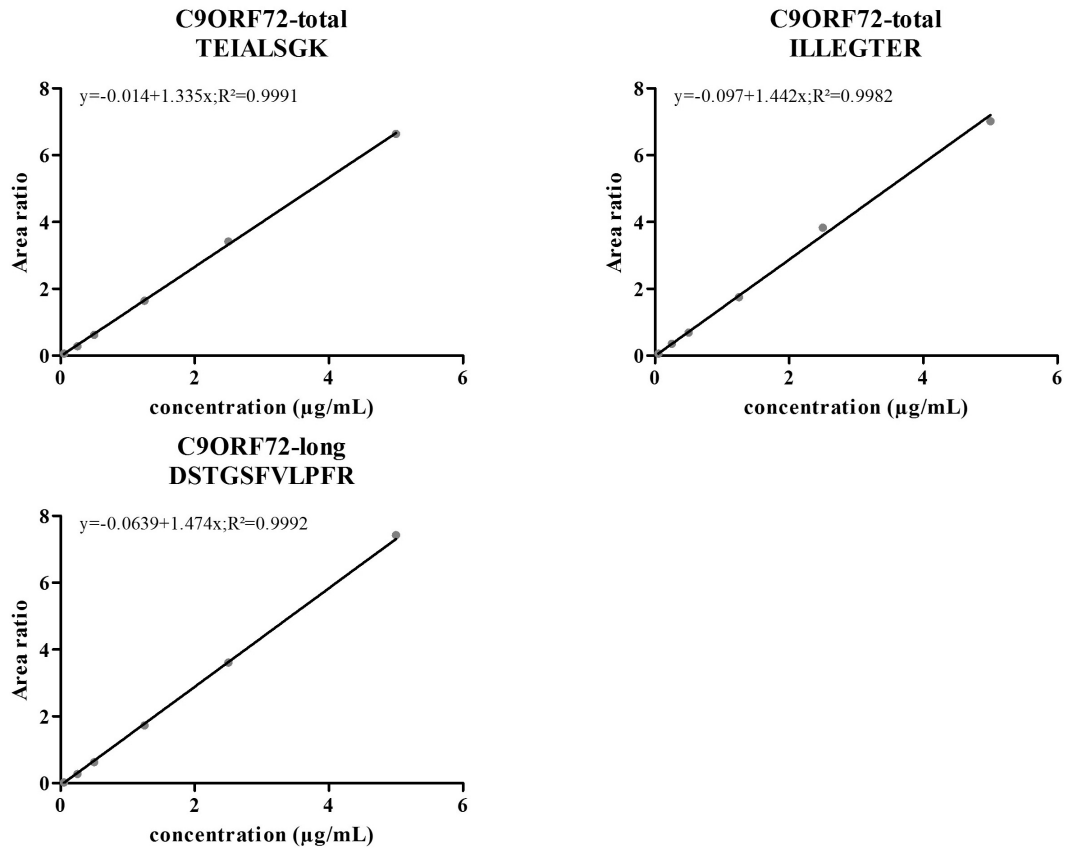

FIGURE 5 | Standard curves of the three selected peptides for quantification of human C9ORF72 in a mice brain lysate. C9ORF72 recombinant long isoform was spiked in increasing concentrations, allowing the generation of a six points standard curves (weighing 1/x).

validation experiments because it contains all three quantitative peptides.

Linearity and sensitivity of the method were evaluated with a 6-point calibration curve of the recombinant C9-L. Mice brain proteins were extracted with the same protocol as for human brain. The method was shown to be linear from 50 to 5,000 ng/mL with a Lower Limit Of Quantification (LLOQ) observed at 50 ng/mL (**Figure 5**). LLOQ was defined based on accuracy between 80 and 120%. To determine a potential matrix effect between mice and human brains, recombinant C9ORF72 was spiked into human brain extract in triplicates at 500 ng/mL and confronted against the calibration curve in mice (**Table 2**). The three peptides displayed good accuracy in the range of 85–115%, which demonstrate the suitability of mice brain extract as a surrogate matrix for C9ORF72 determination in human brain. Intraday repeatability of the analytical method was evaluated in human brain extract. Briefly, proteins from a human brain sample were extracted and divided in five aliquots for protein precipitation, digestion and LC-MS/MS analysis. Precision was observed below 10% for each peptide (n = 5) (**Table 3**). Interday precision was also evaluated by analyzing three different brain extracts on three different days. Variability was found to be acceptable, with CV% between 10 and 26% (**Table 4**). Stability of the C9ORF72 protein in brain extract during the sample preparation is an important parameter for quantification. Brain extracts were either directly processed or left on ice for 90 min, corresponding to the time to process around 100 samples from protein extraction to protein precipitation. Both conditions showed similar area ratio for the three peptides demonstrating stability of C9ORF72 in our conditions (**Figure 6**).

Taking all results together, the protocol was found efficient for determination of C9ORF72 in human brain samples. The new method demonstrated robustness with variability and accuracy below 20%.

C9ORF72 Isoforms Determination in Human Brain Tissue

Levels of C9ORF72 protein were investigated in frontal cortices of FTD patients (with or without ALS) carrying *C9ORF72* expansions (n = 21), patients with non-genetic FTD, pathologically characterized by TDP-43-positive neuronal inclusions (n = 10), and neurologically normal controls (n = 12) determined by pathologists within the network of the NeuroCeb brain bank.

The three peptides selected for quantification were detected above LLOQ in all samples (**Figure 7**). However, no signal was detected for the unique C9-S peptide in any of the

TABLE 2 | Matrix effect between mice and human brain, evaluated for each quantitative peptide.

QC in human brain (n = 3)	TEIALSGK				ILLEGTER				DSTGSFVLPFR			
	Mean measured (μg/mL)	CV %	Theoretical value (μg/mL)	Bias%	Mean Measured (μg/mL)	CV%	Theoretical value (μg/mL)	Bias %	Mean measured (μg/mL)	CV %	Theoretical value (μg/mL)	Bias %
Blank (endogenous)	0.285	6%			0.292	2%			0.348	5%		
Spike (+0.5 μg/mL)	0.767	4%	0.785	2%	0.815	1%	0.792	3%	0.785	3%	0.848	7%

Human brain lysate was spiked with 0.5 μg/mL of C9ORF72 recombinant long isoform and back calculated against a calibration curve prepared in mice brain lysate. Endogenous C9ORF72 concentration was added to the spiked amount for calculations of bias. Bias and CV% are reported.

TABLE 3 | Intraday variability of the measured C9ORF72 concentrations by replicate analysis of a control human brain sample (n = 5).

Intraday validation (n = 5)	TEIALSGK		ILLEGTER		DSTGSFVLPFR	
	Mean (μg/mL)	CV	Mean (μg/mL)	CV	Mean (μg/mL)	CV
Control-1	0.37	4%	0.43	7%	0.414	6%

TABLE 4 | Interday variation of the measured C9ORF72 concentrations in 2 C9ORF72 human brains and one human control brain samples; (n = 3).

Interday variation (n = 3)	TEIALSGK		ILLEGTER		DSTGSFVLPFR	
	Mean (μg/mL)	CV	Mean (μg/mL)	CV	Mean (μg/mL)	CV
Control-2	0.514	18%	0.545	10%	0.563	14%
C9-1	0.344	26%	0.382	12%	0.390	22%
C9-2	0.153	11%	0.182	16%	0.197	10%

brain samples. Furthermore, similar amounts of total C9ORF72 and C9-L were measured, considering assay accuracy and variability of 20%, revealing by difference, the low abundance of C9-S (**Supplementary Table S4**). Differences in measured concentrations between groups were assessed by a Mann–Whitney nonparametetric test. Quantification based on any of the three peptides demonstrated a significant decrease in total C9ORF72 in FTD patient with *C9ORF72* expansions ($p < 0.0001$) compared to controls and non-genetic FTD ($p < 0.05$) (**Figure 7**). The decrease in concentration of about 42%, which was observed for the three peptides (i.e., the two peptides representing total C9ORF72 and the one specific to the long isoform), corroborates prior Western blot findings.

DISCUSSION

C9ORF72 isoforms have solely been investigated by Western blot which relies on the availability of antibodies whose specificity has to be characterized and validated (Liu et al., 2006; Davidson et al., 2018). In this respect, we developed a MS assay avoiding antibodies and the gold standard immunoprecipitation for protein quantification in complex matrixes. An efficient protocol, with optimized sample preparation steps, i.e., protein extractions from tissue, pellet digestion and SPE, was successfully implemented for the first time to quantitate C9ORF72 isoforms level in a cohort of human brain samples from *C9ORF72* or non-genetic FTD patients and control individuals.

Preparation of tissue lysates and protein extraction is a key issue for proteome coverage, especially the more challenging membrane or nuclear proteins (Cox and Emili, 2006; Wiśniewski et al., 2009). Taking into account the reported distinct subcellular localizations of C9ORF72 isoforms with localization of C9-S to the nuclear membrane and C9-L in the cytoplasm (Xiao et al., 2015), lysis protocols were investigated. We found that Triton X-100 facilitated the extraction of the C9ORF72 proteins, in agreement with higher extraction yield reported with detergent based protocols in comparison to organic solvents or chaotropic reagents such as urea, in fatty tissue such as the brain (Shevchenko et al., 2012). In addition, the Triton X-100 protein extraction protocol probably enhanced the subsequent trypsin digestion as previously reported (An et al., 2015). A simple pellet digestion protocol (Becher et al., 2017) was used here for removal of potentially interfering matrix components such as small molecules, phospholipids, peptides (Ouyang et al., 2012) and the added Triton X-100 surfactant which could otherwise have a dramatic impact on MS sensitivity (Cox and Emili, 2006). Considering the molecular weight difference between C9-S and C9-L, i.e., 25 and 54 kDa, and the potential solubility of smaller proteins in organic precipitation solvents (Lopes et al., 2014), we ascertained the equivalent recovery for both isoforms, ensuring accurate quantification. The signal observed for both C9-S and C9-L unique peptides in transfected HEK293 cells further confirmed the correct extraction and precipitation of C9ORF72 isoforms.

The validation experiments demonstrated that the precision provided by the method was satisfactory, with coefficients of variation below 20% and its ability to determine C9ORF72 relative concentration in human brain samples with a LLOQ at 50 ng/mL. Prior work with pellet digestion reported LLOQ around 1 μg/mL for quantification of therapeutic antibodies in plasma (Ouyang et al., 2012; Becher et al., 2017). In the present work, the use of PRM mode for peptide detection offered additional selectivity linked to the high resolution and mass precision measurements in the Orbitrap and the opportunity to accumulate fragment ions in the C-trap, eventually summed-up for higher signal intensity. Regarding signature peptide selection, chimeric or humanized therapeutic antibodies are more and more alike and so only a few peptides are unique

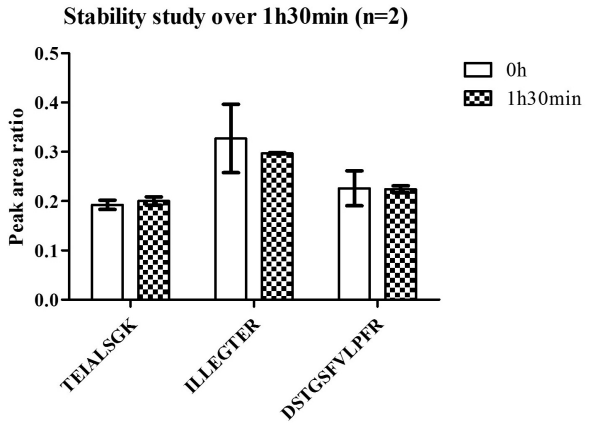

FIGURE 6 | Stability of C9ORF72 in human brain at room temperature.

to one therapeutic antibody. These unique peptides do not necessarily have the best physicochemical properties to be efficiently ionized by electrospray ionization whereas C9ORF72 peptides were selected based on signal intensity. Also, the fivefold lower protein content in brain extract found by BCA may have contributed to reduced deleterious matrix effects compared to plasma. Most probably, a combination of these factors explains the gain in sensibility obtained for C9ORF72.

So far, C9ORF72 quantification was only accomplished by Western blot using in-house generated antibodies. The new MS assay allowed for the first time monitoring of C9ORF72 isoforms in human brain samples, without potential interference due to cross-reactivity of antibodies. First, we observed equivalent brain levels for C9-total and C9-L, indicating a low abundance of C9-S within the assay variability, i.e., a concentration below 20% of C9-L. Here, a similar peptide release between the recombinant and endogenous C9-S/C9-L was assumed since the protocol denature proteins during tissus lysis and pellet digestion which enhances digestion by protein unfolding (Pritchard et al., 2014). The low abundance of C9-S is further strenghtened by the undetected unique C9-S peptide regardless of the evaluated brain protein extraction protocol, including those previously published for C9ORF72 (Xiao et al., 2015), whereas the peptide was well detectable in transfected HEK293 cells. Considering the lower abundance of the short isoform, deeper fractionation of the sample, for instance through enrichment of gray matter

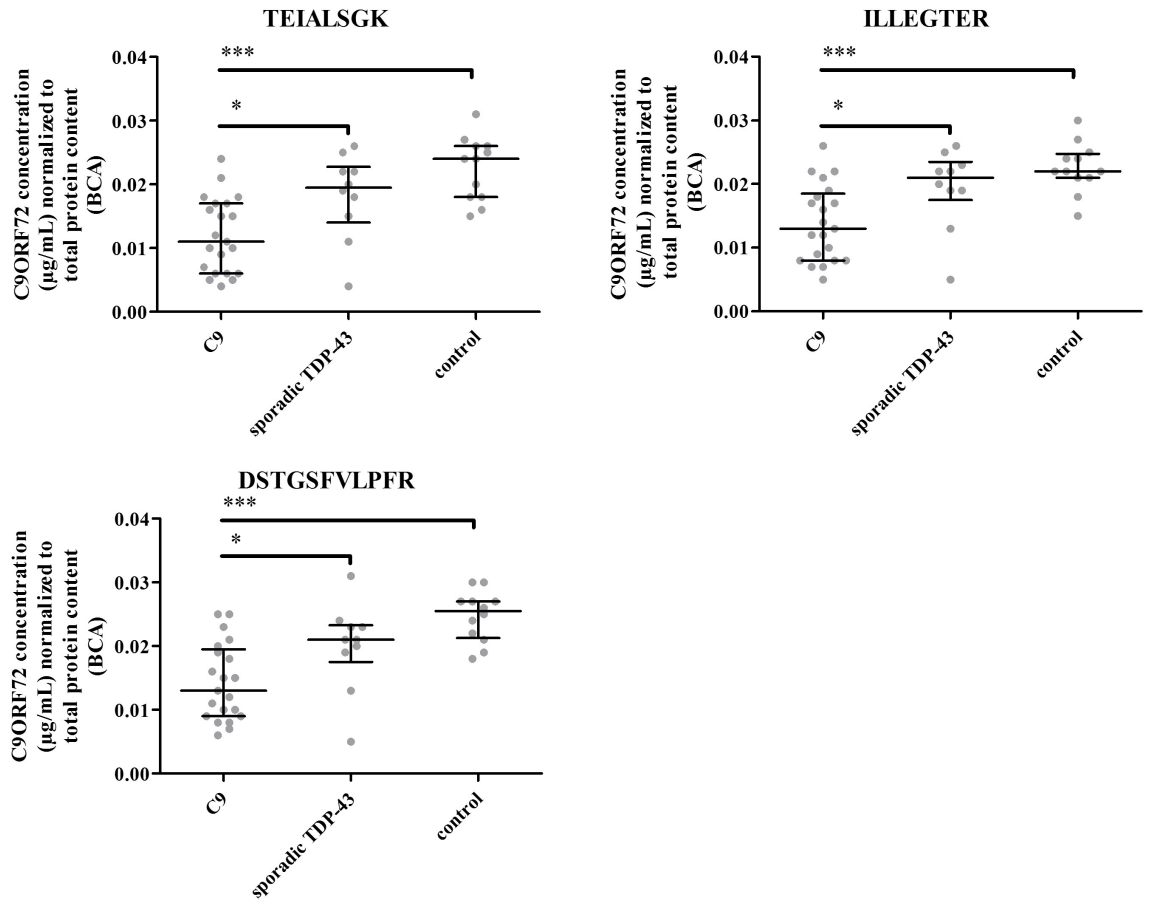

FIGURE 7 | C9ORF72 concentration determination based on peptides 1, 2, and 3 in the human brain frontal cortex cohort. *$P \leq 0.05$ and ***$P \leq 0.001$.

or subcellular fractions, could be considered to increase assay sensitivity. Also, post-translational modifications could impact peptide detection by our targeted method, even though none was reported for any of the selected peptides by now. In previous Western blot reports, Waite et al. (Waite et al., 2014) found that C9-S was in lower abundance than C9-L but questioned the specificity of their antibody, two bands being present at 27 and 29 kDa. Although, Xiao et al. (2015) detected the C9-S in frontal cortex using isoform specific antibodies, the relative abundance between C9-L and C9-S was not determined, probably in relation to the inherent limitation of protein quantification by Western blotting which depends on the affinity and specificity of the reagents (Aebersold et al., 2013). Next, we were able to confirm in a collection of 43 frontal cortices the diminution of the C9ORF72 long protein concentration in *C9ORF72* FTD patients corroborating previous observations by the two Westernblot studies (Waite et al., 2014; Xiao et al., 2015) and at the mRNA level. The age of onset of the disease is highly variable as well as the number of expansion, the precise determination of C9ORF72 levels afforded by the new assay can be used to investigate correlations between the length of expansion and the levels of C9ORF72.

In summary, an efficient protocol was developed for quantification of C9ORF72 isoforms in brain samples by MS. Combination of optimal sample preparation and targeted high-resolution MS demonstrated robust and efficient quantification ability. This new assay has the advantage of being based on MS, avoiding the potential cross-reactivity of antibodies and simplifying implementation in various laboratories (Addona et al., 2009). C9ORF72 long isoform was significantly decreased in carriers of *C9ORF72* expansion in comparison with controls and non-genetic FTD patients with or without ALS, corroborating prior observations made by Western blot and at the mRNA level. Whereas some studies reported a short isoform, here it represents less than 20% of the long one, suggesting possible non-specificity or cross-reactivity of antibodies. To our knowledge, this is the first report of a MS-based quantification assay for C9ORF72 proteins. This method needs to be further applied to other biological matrixes of a more relevant diagnostic nature and potentially to follow treatment efficacy in the future. This method could be easily implemented to mice models of *C9ORF72* FTD or other animal model owing to a highly conserved sequence, in order to advance understanding of the contribution of C9ORF72 to disease mechanisms.

AUTHOR CONTRIBUTIONS

AV, CF, ILB, FL, VA, and FB contributed conception and design of the study. AV, CF, AC, ML, and VA conducted the experiments. NeuroCEB brain Bank provided samples. AV and FB wrote the first draft of the manuscript. ILB, CF, and VA wrote sections of the manuscript. CF, FF, FE, ILB, CJ, and FL critically reviewed the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

The research leading to these results has received funding from the program "Investissements d'avenir" ANR-10-IAIHU-06.

ACKNOWLEDGMENTS

The authors thank the "NeuroCEB" Brain Bank (GIE Neuro-CEB BB-0033-00011); the brain donation program is funded by a consortium of patient associations: ARSEP (association for research on multiple sclerosis), CSC (association for research on cerebellar ataxias), France Parkinson, France Alzheimer and LECMA-Vaincre Alzheimer. The NeuroCEB Brain bank network includes: Prof. Franck Letournel (CHU Angers), Prof. Anne Vital (CHU Bordeaux), Prof. Françoise Chapon (CHU Caen), Dr. Catherine Godfraind (CHU Clermont-Ferrand), Prof. Claude-Alain Maurage (CHU Lille), Dr. Vincent Deramecourt (CHU Lille), Dr. David Meyronnet (CHU Lyon), Dr. Nathalie Streichenberger (CHU Lyon), Dr. André Maues de Paula (CHU Marseille), Dr. Valérie Rigau (CHU Montpellier), Dr. Fanny Vandenbos-Burel (Nice), Prof. Charles Duyckaerts (CHU PS Paris), Prof. Danielle Seilhean (CHU PS, Paris), Dr. Véronique Sazdovitch (CHU PS Paris), Dr. Serge Milin (CHU Poitiers), Dan Christian Chiforeanu (CHU Rennes), Prof. Annie Laquerrière (CHU Rouen), Dr. Béatrice Lannes (CHU Strasbourg).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2018.00589/full#supplementary-material

REFERENCES

Addona, T. A., Abbatiello, S. E., Schilling, B., Skates, S. J., Mani, D. R., Bunk, D. M., et al. (2009). Multi-Site assessment of the precision and reproducibility of multiple reaction monitoring–based measurements of proteins in plasma. *Nat. Biotechnol.* 27, 633–641. doi: 10.1038/nbt.1546

Aebersold, R., Burlingame, A. L., and Bradshaw, R. A. (2013). Western blots versus selected reaction monitoring assays: time to turn the tables? *ASBMB* 12, 2381–2382.

An, B., Zhang, M., Johnson, R. W., and Qu, J. (2015). Surfactant-aided precipitation/on-pellet-digestion (SOD) procedure provides robust and rapid sample preparation for reproducible, accurate and sensitive LC/MS quantification of therapeutic protein in plasma and tissues. *Anal. Chem.* 87, 4023–4029. doi: 10.1021/acs.analchem.5b00350

Becher, F., Ciccolini, J., Imbs, D. C., Marin, C., Fournel, C., Dupuis, C., et al. (2017). A simple and rapid LC-MS/MS method for therapeutic drug monitoring of cetuximab: a GPCO-UNICANCER proof of concept study in head-and-neck cancer patients. *Sci. Rep.* 7:2714. doi: 10.1038/s41598-017-02821-x

Chen, D., Wang, Y., and Chin, E. R. (2015). Activation of the endoplasmic reticulum stress response in skeletal muscle of G93A*SOD1 amyotrophic lateral sclerosis mice. *Front. Cell. Neurosci.* 9:170. doi: 10.3389/fncel.2015.00170

Chen, Y. T., Tuan, L. P., Chen, H. W., Wei, I. A., Chou, M. Y., Chen, H. M., et al. (2015). Quantitative analysis of prostate specific antigen isoforms using immunoprecipitation and stable isotope labeling mass spectrometry. *Anal. Chem.* 87, 545–553. doi: 10.1021/ac5033066

Ciura, S., Lattante, S., Le Ber, I., Latouche, M., Tostivint, H., Brice, A., et al. (2013). Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Ann. Neurol.* 74, 180–187. doi: 10.1002/ana.23946

Cox, B., and Emili, A. (2006). Tissue subcellular fractionation and protein extraction for use in mass-spectrometry-based proteomics. *Nat. Protoc.* 1, 1872–1878. doi: 10.1038/nprot.2006.273

Davidson, Y. S., Robinson, A. C., Rollinson, S., Pickering-Brown, S., Xiao, S., Robertson, J., et al. (2018). Immunohistochemical detection of C9orf72 protein in frontotemporal lobar degeneration and motor neurone disease: patterns of immunostaining and an evaluation of commercial antibodies. *Amyotrophic Lateral Sclerosis Frontotemporal Degener.* 19, 102–111. doi: 10.1080/21678421.2017.1359304

DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-Linked FTD and ALS. *Neuron* 72, 245–256. doi: 10.1016/j.neuron.2011.09.011

Dupré, M., Gilquin, B., Fenaille, F., Feraudet-Tarisse, C., Dano, J., Ferro, M., et al. (2015). Multiplex quantification of protein toxins in human biofluids and food matrices using immunoextraction and high-resolution targeted mass spectrometry. *Anal. Chem.* 87, 8473–8480. doi: 10.1021/acs.analchem.5b01900

Gong, C., Zheng, N., Zeng, J., Aubry, A. F., and Arnold, M. E. (2015). Post-pellet-digestion precipitation and solid phase extraction: a practical and efficient workflow to extract surrogate peptides for ultra-high performance liquid chromatography – tandem mass spectrometry bioanalysis of a therapeutic antibody in the low ng/mL range. *J. Chromatogr. A* 1424, 27–36. doi: 10.1016/j.chroma.2015.10.049

Jedrychowski, M. P., Wrann, C. D., Paulo, J. A., Gerber, K. K., Szpyt, J., Robinson, M. M., et al. (2015). Detection and quantitation of circulating human irisin by tandem mass spectrometry. *Cell Metab.* 22, 734–740. doi: 10.1016/j.cmet.2015.08.001

Lagier-Tourenne, C., Baughn, M., Rigo, F., Sun, S., Liu, P., Li, H. R., et al. (2013). Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc. Natl. Acad. Sci. U.S.A.* 110, E4530–E4539. doi: 10.1073/pnas.1318835110

Lesur, A., Ancheva, L., Kim, Y. J., Berchem, G., van Oostrum, J., and Domon, B. (2015). Screening protein isoforms predictive for cancer using immunoaffinity capture and fast LC-MS in PRM mode. *PROTEOMICS – Clin. Appl.* 9, 695–705. doi: 10.1002/prca.201400158

Levine, T. P., Daniels, R. D., Gatta, A. T., Wong, L. H., and Hayes, M. J. (2013). The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics (Oxford, England)* 29, 499–503. doi: 10.1093/bioinformatics/bts725

Liu, Y. T., Kardosh, A., Cooc, J., and Schönthal, A. H. (2006). Potential misidentification of cyclooxygenase-2 by western blot analysis and prevention through the inclusion of appropriate controls. *Mol. Biotechnol.* 34, 329–335.

Lopes, F., Cowan, D. A., Thevis, M., Thomas, A., and Parkin, M. C. (2014). Quantification of intact human insulin-like growth factor-i in serum by nano-ultrahigh-performance liquid chromatography/tandem mass spectrometry: quantification of insulin-like growth factor-I by nanoUHPLC-MS/MS. *Rapid Commun. Mass Spectrom.* 28, 1426–1432. doi: 10.1002/rcm.6908

MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., et al. (2010). Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26, 966–968. doi: 10.1093/bioinformatics/btq054

Majounie, E., Renton, A. E., Mok, K., Dopper, E. G. P., Waite, A., Rollinson, S., et al. (2012). Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* 11, 323–330. doi: 10.1016/S1474-4422(12)70043-1

Mori, K., Weng, S. M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., et al. (2013). The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* 339, 1335–1338. doi: 10.1126/science.1232927

Ouyang, Z., Furlong, M. T., Wu, S., Sleczka, B., Tamura, J., Wang, H., et al. (2012). Pellet digestion: a simple and efficient sample preparation technique for LC–MS/MS quantification of large therapeutic proteins in plasma. *Bioanalysis* 4, 17–28. doi: 10.4155/bio.11.286

Pritchard, C., Groves, K. J., Biesenbruch, S., O'Connor, G., Ashcroft, A. E., Arsene, C., et al. (2014). Quantification of human growth hormone in serum with a labeled protein as an internal standard: essential considerations. *Anal. Chem.* 86, 6525–6532. doi: 10.1021/ac501032q

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., et al. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain* 134, 2456–2477. doi: 10.1093/brain/awr179

Renton, A. E., Majounie, E., Waite, A., Simón-Sánchez, J., Raphael Gibbs, S. R. J., Schymick, J. C., et al. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72, 257–268. doi: 10.1016/j.neuron.2011.09.010

Rosso, S. M. (2003). Frontotemporal dementia in the netherlands: patient characteristics and prevalence estimates from a population-based study. *Brain* 126, 2016–2022. doi: 10.1093/brain/awg204

Shevchenko, G., Musunuri, S., Wetterhall, M., and Bergquist, J. (2012). Comparison of extraction methods for the comprehensive analysis of mouse brain proteome using shotgun-based mass spectrometry. *J. Proteome Res.* 11, 2441–2451. doi: 10.1021/pr201169q

Torsetnes, S. B., Nordlund, M. S., Paus, E., Halvorsen, T. G., and Reubsaet, L. (2013). Digging deeper into the field of the small cell lung cancer tumor marker ProGRP: a method for differentiation of its isoforms. *J. Proteome Res.* 12, 412–420. doi: 10.1021/pr300751j

Umoh, M. E., Dammer, E. B., Dai, J., Duong, D. M., Lah, J. J., Levey, A. I., et al. (2018). A proteomic network approach across the ALS-FTD disease spectrum resolves clinical phenotypes and genetic vulnerability in human brain. *EMBO Mol. Med.* 10, 48–62. doi: 10.15252/emmm.201708202

Waite, A. J., Bäumer, D., East, S., Neal, J., Morris, H. R., Ansorge, O., et al. (2014). Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiol. Aging* 35, 1779.e5–1779.e13. doi: 10.1016/j.neurobiolaging.2014.01.016

Wiśniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nat. Methods* 6, 359–362. doi: 10.1038/nmeth.1322

Xiao, S., MacNair, L., McGoldrick, P., McKeever, P. M., McLean, J. R., Zhang, M., et al. (2015). Isoform-specific antibodies reveal distinct subcellular localizations of C9orf72 in amyotrophic lateral sclerosis: C9orf72 isoforms in ALS. *Ann. Neurol.* 78, 568–583. doi: 10.1002/ana.24469

Xiao, S., MacNair, L., McLean, J., McGoldrick, P., McKeever, P., Soleimani, S., et al. (2016). C9orf72 isoforms in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Brain Res.* 1647, 43–49. doi: 10.1016/j.brainres.2016.04.062

Zhang, D., Iyer, L. M., He, F., and Aravind, L. (2012). Discovery of novel DENN proteins: implications for the evolution of eukaryotic intracellular membrane structures and human disease. *Front. Genet.* 3:283. doi: 10.3389/fgene.2012.00283

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a shared affiliation, though no other collaboration, with one of the authors FE.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 05 December 2018
doi: 10.3389/fnins.2018.00896

Microbiome-Derived Lipopolysaccharide (LPS) Selectively Inhibits Neurofilament Light Chain (NF-L) Gene Expression in Human Neuronal-Glial (HNG) Cells in Primary Culture

Walter J. Lukiw[1,2,3], Lin Cong[4], Vivian Jaber[1] and Yuhai Zhao[1,5]*

[1] Neuroscience Center, Louisiana State University School of Medicine, Louisiana State University Health Sciences Center, New Orleans, LA, United States, [2] Department of Neurology, Louisiana State University School of Medicine, Louisiana State University Health Sciences Center, New Orleans, LA, United States, [3] Department of Ophthalmology, Louisiana State University School of Medicine, Louisiana State University Health Sciences Center, New Orleans, LA, United States, [4] Department of Neurology, Shengjing Hospital, China Medical University, Shenyang, China, [5] Department of Anatomy and Cell Biology, Louisiana State University School of Medicine, Louisiana State University Health Sciences Center, New Orleans, LA, United States

OPEN ACCESS

Edited by:
Ioannis Dragatsis, University of Tennessee Health Science Center (UTHSC), United States

Reviewed by:
Fernanda Marques, Instituto de Pesquisa em Ciências da Vida e da Saúde (ICVS), Portugal
Cheryl Wellington, University of British Columbia, Canada

***Correspondence:**
Walter J. Lukiw
wlukiw@lsuhsc.edu

Specialty section:
This article was submitted to Neurogenomics, a section of the journal Frontiers in Neuroscience

Received: *07 March 2018*
Accepted: *16 November 2018*
Published: *05 December 2018*

Citation:
Lukiw WJ, Cong L, Jaber V and Zhao Y (2018) Microbiome-Derived Lipopolysaccharide (LPS) Selectively Inhibits Neurofilament Light Chain (NF-L) Gene Expression in Human Neuronal-Glial (HNG) Cells in Primary Culture. Front. Neurosci. 12:896. doi: 10.3389/fnins.2018.00896

The remarkable co-localization of highly pro-inflammatory lipopolysaccharide (LPS) with sporadic Alzheimer's disease (AD)-affected neuronal nuclei suggests that there may be some novel pathogenic contribution of this heat stable neurotoxin to neuronal activity and neuron-specific gene expression. In this communication we show for the first time: (i) the association and envelopment of sporadic AD neuronal nuclei with LPS in multiple AD neocortical tissue samples; and (ii) a selective repression in the output of neuron-specific neurofilament light (NF-L) chain messenger RNA (mRNA), perhaps as a consequence of this association. The down-regulation of NF-L mRNA and protein is a characteristic attribute of AD brain and accompanies neuronal atrophy and an associated loss of neuronal architecture with synaptic deficits. To study this phenomenon further, human neuronal-glial (HNG) cells in primary culture were incubated with LPS, and DNA arrays, Northern, Western, and ELISA analyses were used to quantify transcription patterns for the three member neuron-specific intermediate filament-gene family NF-H, NF-M, and NF-L. As in sporadic AD limbic-regions, down-regulated transcription products for the NF-L intermediate filament protein was significant. These results support our novel hypothesis: (i) that internally sourced, microbiome-derived neurotoxins such as LPS contribute to a progressive disruption in the read-out of neuron-specific genetic-information; (ii) that the presence of LPS-enveloped neuronal nuclei is associated with a down-regulation in NF-L expression, a key neuron-specific cytoskeletal component; and (iii) this may have a bearing on progressive neuronal atrophy, loss of synaptic-contact and disruption of neuronal architecture, all of which are characteristic pathological

features of sporadic-AD brain. This is the first report that provides evidence for a neuron-specific effect of a human GI-tract microbiome-derived neurotoxin on decreased NF-L abundance in both sporadic AD temporal lobe neocortex *in vivo* and in LPS-stressed HNG cells *in vitro*.

Keywords: Alzheimer's disease (AD), DNA array, inflammatory degeneration, lipopolysaccharide (LPS), microbiome, neurofilament light chain (NF-L), neurofilament triplet

INTRODUCTION

Recently there has been a resurgence of interest in the human gastrointestinal (GI) tract microbiome and its potential contribution to human health and disease. One area receiving considerable research attention has been the possible involvement of human GI-tract microbiome-derived neurotoxins with progressive and terminal neurological diseases associated with aging and inflammatory neurodegeneration. These microbiome-derived neurotoxic exudates consist of a remarkably complex and neurobiologically potent array of pro-inflammatory endotoxins and exotoxins (such as fagilysin), lipooligosaccharides (LOS), lipopolysaccharides (LPS; including the extremely pro-inflammatory *B. fragilis* LPS, BF-LPS), microRNA-like small non-coding RNAs (sncRNA), and an extensive variety of bacterial-derived amyloids (Hofer, 2014; Foster et al., 2016; Lukiw, 2016a,b; Zhan et al., 2016, 2018; Mancuso and Santangelo, 2017; Yang and Chiu, 2017; Zhao et al., 2017a,b,c; Zhao and Lukiw, 2018). Several recent papers have addressed the emerging link between elements of the human GI-tract microbiome and Alzheimer's disease (AD), a common, chronic, and progressive age-related neurodegenerative disease whose incidence is reaching epidemic proportions and represents a major, lethal, neuropsychiatric disorder that currently constitutes a major healthcare concern worldwide (Zhan et al., 2016; Jiang et al., 2017; Cox and Weiner, 2018; Szablewski, 2018; Zhao and Lukiw, 2018).

Both the familial and the much more common sporadic forms of AD are characterized by the appearance of extracellular deposits including dense, insoluble amyloid-beta (Aβ) peptide enriched senile plaques (SP) and tau- and neurofilament-protein enriched neurofibrillary tangles (NFT), and neuropathologically by the progressive atrophy of large neurons, ensuing loss of synaptic contacts and altered neuronal cytoarchitecture (Clark and Vissel, 2015; Zhao and Lukiw, 2015; Minter et al., 2016). We adopted the strategy that because inter-synaptic connections, the radial diameter of neurons and the overall neuronal architecture and morphology are maintained in large part by this relatively abundant three member neuron-specific neurofilament gene family – encoding the neurofilament light (NF-L; NEFL; 68 kDa), neurofilament medium (NF-M; ∼160 kDa), and neurofilament heavy (NF-H; ∼205 kDa) chain proteins – we reasoned that NF-L, NF-M, and NF-H relative abundance would be an experimentally practical and suitable choice to look for changes in expression in both sporadic AD brain and in LPS treated human neuronal-glial (HNG) cells in primary culture.

Our findings indicate for the first time, that linked to a progressive association of the amphiphilic glycolipid LPS with sporadic AD neuronal nuclei there appears to be a parallel and selective repression in the output of neuron-specific NF-L mRNA in AD brain compared to age-and gender-matched controls. This is noteworthy because down-regulation of NF-L expression is a characteristic feature of the limbic system in AD and accompanies the atrophy of neurons and progressive loss of neuronal architecture and synaptic contact in the AD brain (McLachlan et al., 1988; Lukiw et al., 1990; Julien and Mushynski, 1998; Clement et al., 2016; Khalil et al., 2018). These effects were further observed in HNG primary cultures incubated with Gram-negative bacterial-derived LPS in which was observed a significant LPS-mediated down-regulation of the NF-L intermediate filament protein. Taken together these results suggest: (i) that microbiome-derived LPS may contribute to a progressive disruption in the read-out of the brain's neuron-specific genetic-information; (ii) that NF-L mRNA and the expression of NF-L proteins are one important neuron-specific transcript targeted by microbiome-derived LPS; and (iii) that this may have a bearing on neuronal atrophy, disruption of the neuronal architecture and loss of synaptic organization, all of which are characteristic neuropathological features of AD-affected brain.

MATERIALS AND METHODS

Human Brain Tissues, Antibodies and Immunohistochemistry

Female control [N = 12; mean age ± one standard deviation of 85.8 ± 2.1 years with a post-mortem interval (PMI) of (mean ± one standard deviation) 3.6 ± 1.5 h] and age-matched AD (N = 12; age 87.7 ± 2.5 years and PMI 3.8 ± 1.2 h) human superior temporal lobe neocortical tissues (Brodmann A22) were obtained from the University of Maryland, from archived material at the University of Toronto and the Louisiana State University Neuroscience Center, and the University of California (UC)-Irvine Brain Bank. A total of 24 female, age-, gender-, and PMI-matched control and AD brains were examined for LPS content. For immunocytochemistry of LPS human brain tissue samples were embedded in OCT and frozen at −80°C; 10 μm brain sections were cut using a Shandon cryotome (Waltham, MA, United States). After an initial fixation with 4% paraformaldehyde for 20 min, sections were then incubated in primary antibodies (1:1000; 1 × PBS with 2% BSA, 2% goat or donkey serum and 0.1% TX-100) overnight at 4°C, washed with PBS, and then incubated with Alexa Fluor-conjugated species-specific secondary antibodies (ThermoFisher Scientific,

Waltham, MA, United States) for 3 h at RT (see further details below). Sections were counter-stained with DAPI for nuclei, followed by quenching with Autofluorescence Eliminator Reagent (Millipore Cat No. 2160; Zhan et al., 2016; Zhao et al., 2017a,b,c), mounted on glass slides, cover-slipped with Fluoromount-G (ThermoFisher Scientific) and imaged using a Zeiss LSM 700 Confocal Laser Scanning microscope system (Carl Zeiss Microscopy, Thornwood, NY, United States; Bagyinszky et al., 2017; Zhao et al., 2017a,b,c[1]).

Human Neuronal-Glial (HNG) Cells in Primary Co-culture

Human neuronal-glial primary cells, cryopreserved at first passage one, were obtained from commercial sources and cultured according to the manufacturer's instructions (Lonza PT-2599, Lonza Cell Systems, Allendale, NJ, United States or Cell Systems, ACBRI 376, Kirkland, WA, United States). HNG cells tested negative for HIV-1, HBV, HCV, mycoplasma, bacteria, yeast, and fungi at source, and have been extensively used for studies on brain gene induction and gene expression, neuronal development, neurotoxicology, neuropharmacology, and in *in vitro* models of AD and other age-related neurological disorders that exhibit a progressive age-related inflammatory neurodegeneration (Li et al., 2011; Zhao et al., 2017a,b,c; Zhao and Lukiw, 2018). HNG cells demonstrate particular neuronal and astroglial cell markers including neuron-specific β-tubulin III (βtubIII; red staining; λ_{max} = 690 nm) and glial fibrillary acidic protein (GFAP; glial-specific green stain; λ_{max} = 520 nm). Briefly, HNG cells were maintained as free-floating aggregates (neurospheres) in 75 cm^2 uncoated plastic flask in neural progenitor maintenance media (NPMM; Lonza CC-3209) supplemented with recombinant human fibroblast growth factor (rhFGF) and epidermal growth factor [rhEGF]) and neural survival factor-1 [NSF-1] (Lonza CC-4242) and gentamicin/amphotericin-B (Lonza GA-1000). Differentiation was induced by plating neurospheres onto eight-well glass chamber-slides pre-coated with poly-L-ornithine (an amino acid polymer used as substratum to improve neuronal adhesion); cells were kept at 37°C in a humidified 5% CO_2 atmosphere incubator at all times. The differentiation media (Lonza CC-4242) was free of growth factors but contained NSF and gentamicin/amphotericin-B, 25 ng/ml of brain-derived neurotrophic factor (BDNF), and 1% of fetal bovine serum (FBS). Upon deprivation of growth factors neurospheres began to attach to the well bottoms and next migrated out to form a co-culture of human neurons and glial cells (HNG). HNG cells were used 2 weeks after induction of differentiation; HNG cells initially contained about 5 × 10^5 cells/ml volume and were cultured to ∼70% confluency in HNG cell medium as described in detail (Cui et al., 2010; Bhattacharjee and Lukiw, 2013; Zhao et al., 2014, 2017a,b,c; Lukiw, 2016a,b). HNG cells were subsequently incubated with LPS; the concentration of LPS (Sigma L2630[2]) provided to 2-week-old HNG cells cultured in HNG cell medium was 50 nM for 48 h (see **Figure 4**; for further specific details see also Zhao et al., 2017a,b,c). Higher doses of LPS (up to 5 μM) in HNG cell medium for shorter periods gave comparable results (data not shown).

Immunofluorescence Protocol

Two week old cultures of HNG cells in eight-well chamber slides (BD Biosciences, San Jose, CA, United States) were fixed with 4% paraformaldehyde, then permeabilized and blocked with 0.125% Triton X-100 and 2% normal goat serum in PBS at RT for 1 h. Cells were incubated overnight at 4°C

[1]https://www.mikroskop.com.pl/pdf/LSM700_1.pdf

[2]https://www.sigmaaldrich.com

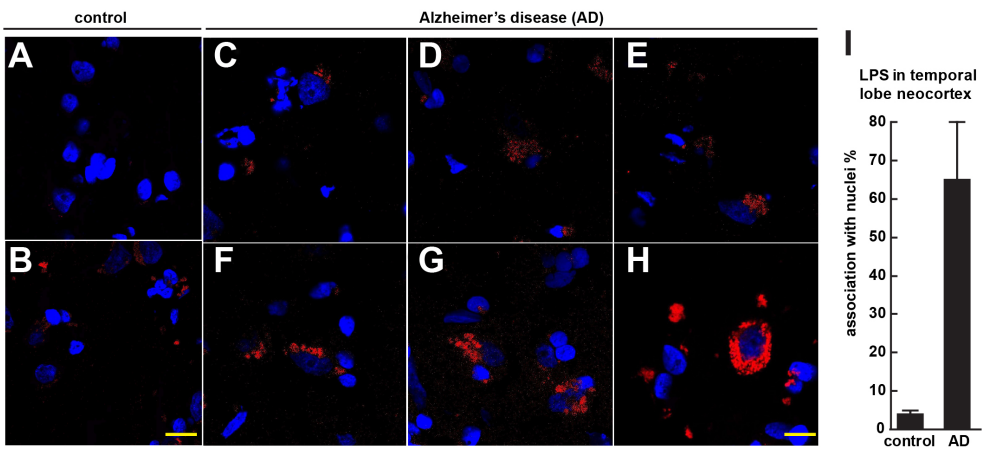

FIGURE 1 | Controls **(A,B)** shows staining of human control superior temporal lobe neocortex (CDR = 0) with LPS (red stain; λ_{max} = 690 nm) and DAPI (blue stain; λ_{max} = 470 nm); Alzheimer's disease **(C–E)** CDR = 2.0 and **(F–H)** CDR = 3.0 illustrates increasing presence of LPS and association of superior temporal lobe AD neocortical nuclei with LPS; LPS staining was observed in controls but the staining was relatively rare and punctate compared to LPS in AD **(C–H)** with the observation of the eventual encapsulation of nuclei by LPS **(F)**; (see text); the bar graph in **(I)** quantifies LPS association with nuclei in the temporal lobe neocortex (see text and Zhao et al., 2017a,b); to ascertain whether LPS-associated and LPS-encapsulated AD nuclei were of neuronal or astroglial origin we next counterstained with the neuron-specific green stain NeuN (λ_{max} = 520 nm; see **Figure 2**); magnification 63 × ; scale bar = 20 μm.

with antibodies for β-tubulin III (for neurons; Sigma T8578, Sigma-Aldrich St. Louis, MO, United States) and GFAP (for astrocytes; Sigma G9629). Cells were subsequently washed for three times with PBS and then incubated for 3 h at room temperature with secondary antibodies conjugated with cy3 or FITC fluorescein (Thermofisher A21422 and A11008; ThermoFisher Scientific, Waltham, MA, United States). After washing and drying, slides were applied with mounting medium containing DAPI (1:10,000; Vector Laboratories, Burlingame, CA, United States) and observed under Zeiss Axioplan Inverted Deconvolution Fluorescent Microscope (63× oil immersion lens; Carl Zeiss, Oberkochen, Germany). Positively stained cells were quantified manually using the manual counter function of ImageJ software (NIH). Negative control with quenching was performed as previously reported in detail (Zhao et al., 2017a,b,c); quantification of LPS was analyzed (i) as a percentage of neuronal area (see below); and/or (ii) by counting multiple microscope fields for the quantity of LPS signals (red stain; λ_{max} = 690 nm) associated with DAPI (blue nuclear stain; λ_{max} = 470 nm) (see **Figures 1**–**3**).

Antibodies – Specificity and Validation

We used mouse anti-*E. coli* LPS (Abcam Cat No. ab35654; Cambridge, MA, United States); rabbit anti-NeuN (Cell Signaling, Cat No. 24307), rabbit anti-GFAP (Sigma-Aldrich, Cat No. G4564) (Cui et al., 2010; Zhao et al., 2014; Lukiw, 2016a, 2017; Zhan et al., 2016). LPS antibody specificity and validation was confirmed (i) using Western immunoblot analysis (see **Figure 1** in Zhao et al., 2017a) which corresponded to the manufacturer's published specifications[3]; and (ii) an antibody neutralization/LPS quenching control assay (Zhao et al., 2017a,b,c). We used sandwich ELISA and Western analysis for NF-L protein determination in LPS-treated HNG cells using Abbexa (abx250460; Cambridge, United Kingdom) and/or LifeSpan BioSciences (LSBio; LS-F6701; Seattle, WA, United States) ELISA systems and NF-L (NEFL) monoclonal antibody (DA2; ThermoFisher Scientific, Cat No. MA1-2010) and a beta actin (β-actin) loading control monoclonal antibody (BA3R; ThermoFisher Scientific, Cat No. MA5-15739) and standard Western analysis as has been previously described by our laboratory (Zhao et al., 2017a,c). To ascertain the association of LPS with neuronal cells confocal images of LPS and NeuN staining were imported into ImageJ[4]; RGB images were first converted into images of separate channels (red for LPS; green for NeuN; and blue for DAPI-stained nuclei). A co-localization finder plugin was run to generate images of co-localization of both channels; each co-localization image was converted into an 8-bit image and inverted. Global thresholding was utilized and the cutoff value was adjusted to the point that only highlighted co-localized particles are black on the image against a white background. Particle analysis was next performed to calculate the area size of the co-localization; this value was then divided by the area size of the NeuN or DAPI staining as the percentage of cell area (Zhao et al., 2017a,b).

[3]http://www.abcam.com/e-coli-lps-antibody-ab211144.html

[4]https://imagej.net/Fiji/Downloads

RNA Isolation and Purification, DNA Array, Northern, ELISA, and Western Analysis

Ultrapure chemicals and reagents of the highest grades commercially available were used throughout these experiments. Typically, 10 mM phenylmethylsulfonyl fluoride (PMSF; Sigma) and 1 U human placenta ribonuclease inhibitor (RNasin; Promega Corporation, Madison WI, United States) were employed in the extraction medium to inhibit protease and specific ribonuclease activities in homogenized human brain tissues or HNG cells. Total cellular RNA was isolated using TRIzol Reagent (Invitrogen-ThermoFisher Scientific; Cat No. 15596026) and quality controlled using analysis using and Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara CA, United States). For Northern blots ∼15 μg of total RNA was separated at 4°C on 1.5% agarose/2.2 M formaldehyde gels at 60 V for 15 h with recirculating 20 mM sodium phosphate buffer, pH 7.0. Gels were stained for 10 min with acridine orange, visualized at 340 nm on a UV trans-illuminator and total RNA was blotted onto Biotrans 0.2 μm Biotrans nylon membrane (Cat No. 01811300, MP Biomedicals). NF-H, NF-M, and NF-L probes were prepared to specific activities of 10^8 dpm ^{32}p-labeled dCTP per/μg of DNA as previously described in detail (Clement et al., 2016). Membranes were pre-hybridized for 12 h at 42°C in 50% formamide, 5× Denhardt's solution (containing 0.1% each of Ficoll 400, polyvinylpyrrolidone, and bovine serum albumin), 5× SSC (standard saline citrate, containing 150 mM sodium chloride, 15 mM sodium citrate, pH 7.0), 50 mM sodium phosphate buffer, pH 6.5, 0.1% sodium dodecyl sulfate (SDS), and 350 μg/ml sonicated herring sperm. This was replaced with fresh hybridization solution containing approximately 5×10^7 cpm of heat-denatured cDNA probes and hybridization occurred at 42°C for 30 h. Nylon membranes were washed under conditions of high stringency (two 30-min washes at 2 × SSC/0.1% SDS at room temperature; two 60-min washes at 0.1× SSC/0.5% SDS at 60–65°C and finally two 15-min washes at 1 × SSC/0.1% SDS at room temperature). When used Fuji RX film was exposed at −70°C for 18–72 h using standard autoradiographic imaging techniques; alternately hybridization signals were quantified using a Typhoon FLA 9500 Biomolecular Imager (GE Healthcare). DNA array analysis was performed as extensively described by our group (Colangelo et al., 2002; Clement et al., 2016; Jaber et al., 2017). Sandwich ELISA and/or Western analysis was used for NF-L protein abundance analysis according to the manufacturers' instructions to quantify both NF-L and β-actin control abundance levels (**Figure 4**).

Statistical Analysis, Integrated Bioinformatics Analysis, and Data Interpretation

For NF-H, NF-M, NF-L, β-actin, and GAPDH mRNA abundance analysis all statistical procedures were analyzed using *p*, analysis of variance (ANOVA) a two-way factorial analysis of variance

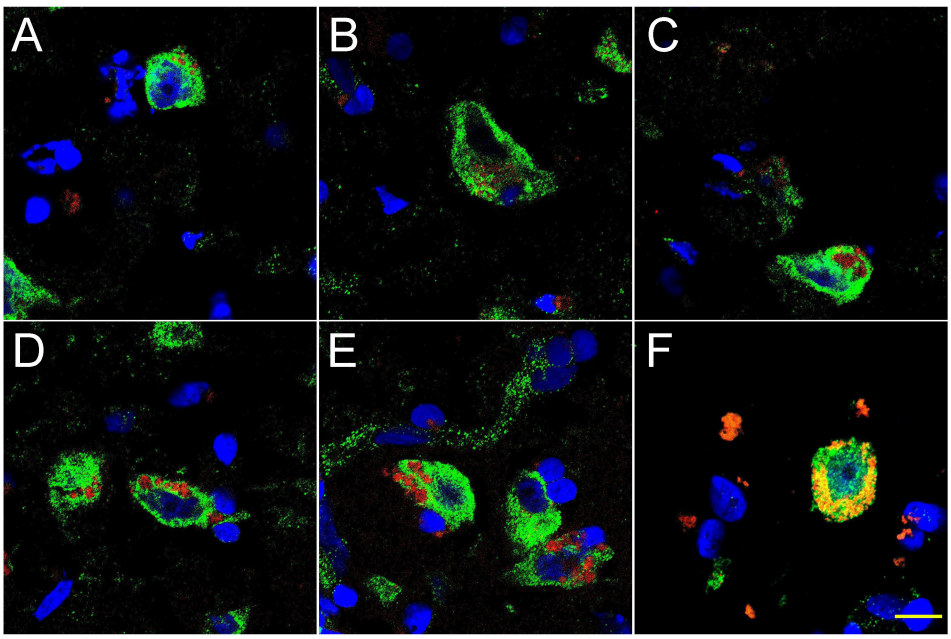

FIGURE 2 | Progressive association and envelopment of AD-affected neocortical neuronal nuclei by LPS (red stain; λ_{max} = 690 nm), DAPI (blue nuclear stain; λ_{max} = 470 nm) and NeuN (neuron-specific green stain; λ_{max} = 520 nm); human superior temporal lobe AD neocortex (Brodmann A22) from CDR (clinical dementia rating) 1.0, 2.0 and 3.0 AD brains; **A,B** = CDR 1.0; **C,D** = CDR 2.0; **E,F** = CDR 3.0 (see also https://knightadrc.wustl.edu/cdr/cdr.htm); LPS staining (red) was subjected to co-localization analysis with the neuronal marker NeuN (green) and/or nuclear marker (blue); magnification 63×; scale bar = 20 μm.

using algorithms, and/or procedures in the SAS language (Statistical Analysis Institute, Cary, NC, United States) and as previously described (Cui et al., 2010; Zhao et al., 2011; Clement et al., 2016; Dendooven and Luisi, 2017; Zhao and Lukiw, 2018). In the results *p*-values of less than 0.05 (ANOVA) were considered to be statistically significant. All NF-H, NF-M, NF-L, β-actin, and GAPDH mRNA abundance data were collected and analyzed using Excel 2016 (Office 365) algorithms (Microsoft Corporation, Redmond WA, United States); all figures were generated using Adobe Illustrator CC 2015 and Photoshop CC version14.0 (Adobe Corporation, San Jose, CA, United States).

RESULTS

Figure 1 (control A,B) shows staining of 10 μm brain sections of human control superior temporal lobe neocortex (Brodmann A22; CDR = 0) of **Figure 1A** an 86 year-old female, 3.6 h PMI and **Figure 1B** an 85-year-old female, 4.3 h PMI) stained with LPS (red stain; λ_{max} = 690 nm) and DAPI (blue stain; λ_{max} = 470 nm); control samples (A) and (B) were from patients with no history of neurodegenerative disease or cognitive impairment; LPS staining was infrequently observed in controls and the staining was sparse; Alzheimer's disease (**Figures 1C–H**); (**Figures 1C–E**) CDR = 2.0; and (**Figures 1F–H**) CDR = 3.0; all female; age-range 79–89; PMI range 3.3–4.1 h; illustrates progressive association and envelopment of AD neocortical nuclei with LPS–LPS (red stain; λ_{max} = 690 nm) and DAPI (blue stain; λ_{max} = 470 nm) staining of human superior temporal lobe AD neocortex (Brodmann A22); note minor 'punctate' LPS signals in control (**Figures 1A,B**) not associated with nuclei, compared to LPS in AD (**Figures 1C–F**) and encapsulation of nuclei by LPS. In a previous study LPS was reported to range from a ∼7- to ∼21-fold increase in abundance in AD brain over age-matched controls from the same anatomical region (Zhao et al., 2017a). The bar graph in **Figure 1I** quantifies LPS association with nuclei in the temporal lobe neocortex (see text); in AD over control; this ratio is about 16.3 for the samples examined; to investigate whether LPS-encapsulated AD nuclei were of neuronal or astroglial origin we next counterstained with the neuron-specific green stain NeuN (**Figure 2**).

Figure 2 shows 10 μM sections of AD neocortical brain tissue stained additionally with NeuN, a neuron-specific green stain (λ_{max} = 520 nm). Interestingly, (i) all LPS staining appears to be confined to one specific nuclear region of AD neuronal nuclei; and (ii) the entire neuronal perinuclear region was occupied by LPS stain in about 5–15% of all neuronal nuclei associated with LPS especially in the later stages of sporadic AD (CDR = 3.0).

Figure 3 shows detail of this LPS-AD neuronal nuclear interaction; approximately 60–70% of all moderate-to-late-stage (CDR 2.0 to 3.0) AD neocortical neuronal nuclei exhibited an association with LPS and about 5–15% of all AD neuronal nuclei in the temporal lobe neocortex showed a complete envelopment by LPS in the 12 sporadic AD cases investigate in this study. Immunocytochemistry further indicated that the "thickness" of "perinuclear LPS envelopes" ranged between 2 and 20 μm (**Figures 2**, **3**).

Figure 4 describes experiments in human neuronal-glial (HNG) cells in primary culture exposed to LPS; HNG cells

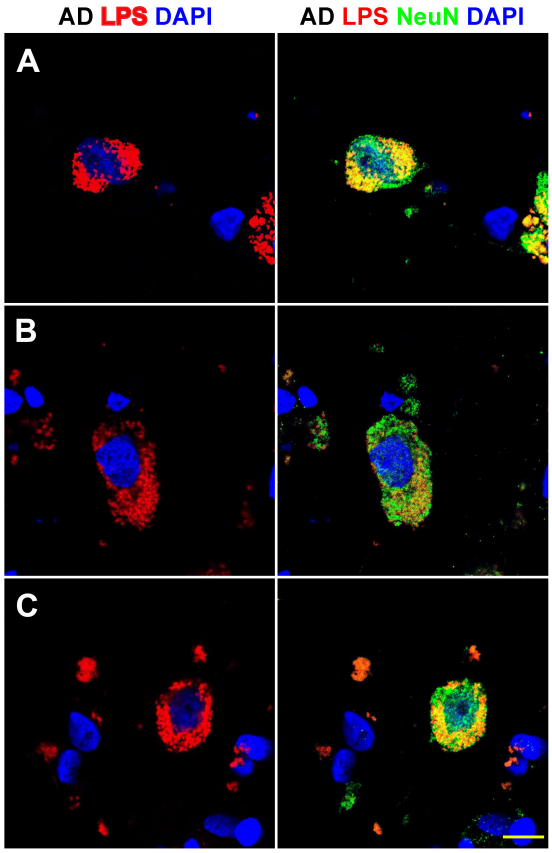

FIGURE 3 | Details of the perinuclear association and envelopment of multiple neuronal nuclei by LPS in sporadic Alzheimer's disease (AD) brain (Brodmann Area 22; superior temporal lobe neocortex); **(A–C)** details of LPS association with neuronal nuclei in three AD independent brain samples; all three AD brain samples were from moderate-to-advanced-AD (CDR 2.0-3.0); no such extensive association between LPS and control brain nuclei was observed (**Figure 1**); lipopolysaccharide (LPS; red stain; λmax = 690 nm); DAPI (nuclear-specific blue stain; λmax = 470 nm); NeuN (neuron-specific green stain; λmax = 520 nm); brain sections are 10 μm thick (see text); a large number of neuronal nuclei in sporadic AD were found to be completely enveloped by LPS especially in the later stages of sporadic AD (CDR 3.0); scale bar for all photos here (lower right) = 20 μm.

(**Figure 4A**) cultured for 2 weeks, 60% confluent and containing about 70% neurons and 30% astroglia, were exposed to 50 nM LPS for 48 h, total RNA was isolated and analyzed on DNA arrays as extensively described by our laboratories (Colangelo et al., 2002; Clement et al., 2016; Jaber et al., 2017). Even at brief periods of exposure to LPS (48 h), LPS appeared to have a strong affinity for DAPI-stained neocortical nuclei (**Figure 4B**). DNA array analysis (**Figure 4C**) indicated that in comparison to the DNA array control transcripts β-actin and glyceraldehyde phosphate dehydrogenase (GAPDH), NF-H mRNA, NF-M mRNA, and NF-L mRNA were found to be reduced to 0.95- 0.72-, and 0.25-fold of controls; in these experiments the NF-L mRNA reduced to 0.25-fold of controls achieved the highest significance of down-regulation ($p < 0.01$, ANOVA); **Figure 4D** represents the quantification of these signals in bar-graph format. **Figure 4E** shows the results of a Northern blot indicating decreased abundance of NF-L mRNA in AD versus age- and gender-matched controls and **Figure 4F** shows the quantified results in bar graph format comparing both the 4.3 and 2.6 knt NF-L mRNA abundance in control and AD. For additional details, please refer to the legend to **Figure 4**.

Figure 5A shows the results of a sandwich ELISA analysis confirming decreased abundance of NF-L protein in AD versus age- and gender-matched controls to about 0.3-fold of control. **Figure 5B** shows the results of a Western analysis of total NF-L protein (MW ∼68 kDa) in control and AD neocortex (left panel) and control and LPS-treated HNG cells (right panel) using β-actin (MW ∼42 kDa) as an internal control and gel-loading marker. **Figure 5C** shows the quantified results in bar graph format comparing NF-L protein abundance in AD and in age- and gender-matched control and in control and LPS-treated HNG cells. Therefore, the independent techniques of ELISA and Western analysis corroborated the observation that the normally highly abundant NF-L protein is reduced in both LPS-enriched AD-affected brain *in vivo* and in LPS-treated HNG cells *in vitro*. For additional experimental details please refer to the legend of **Figure 5**.

DISCUSSION

Gastrointestinal (GI) Tract Microbiome-Derived Neurotoxins in Neurological Disease

The human GI tract, containing about 95% of the entire human microbiome, is the largest reservoir of microorganisms in the human body (Bhattacharjee and Lukiw, 2013; Zhao and Lukiw, 2015; Zhan et al., 2016; Vogt et al., 2017; Zhao et al., 2017a; Zhao and Lukiw, 2018). Comprised of a genetically diverse and densely packed repository of about 100 trillion microorganisms, the GI-tract microbiome is made up of mostly of anaerobic bacterial species with archaebacteria, fungi, protozoa, viruses, and other microbes making up the remainder (Bhattacharjee and Lukiw, 2013; Köhler et al., 2016; Vogt et al., 2017; Zhao et al., 2017a,b,c; Zhao and Lukiw, 2018). Microbial abundance, complexity and speciation, their biophysics, microbiology and neurobiology, molecular genetics, epigenetics, the signaling mechanisms, and pathways involved in microbiome-host communications and interactions are becoming increasingly understood in context of their dynamic contribution to human neurobiology in health, aging, and disease (Bhattacharjee and Lukiw, 2013; Foster et al., 2016; Köhler et al., 2016; Jiang et al., 2017; Cox and Weiner, 2018).

GI-tract derived neurotoxins include an extraordinarily complex mixture of potentially pathogenic amyloid, exotoxins and endotoxins, lipooligosaccahrides (LOS), LPS, and small miRNA-like non-coding RNA (sncRNA) exudates. Normally confined within the healthy human GI-tract, accompanying aging, and disease these secreted neurotoxins can transverse normally protective biophysical and physiological barriers resulting in a persistent systemic inflammatory condition

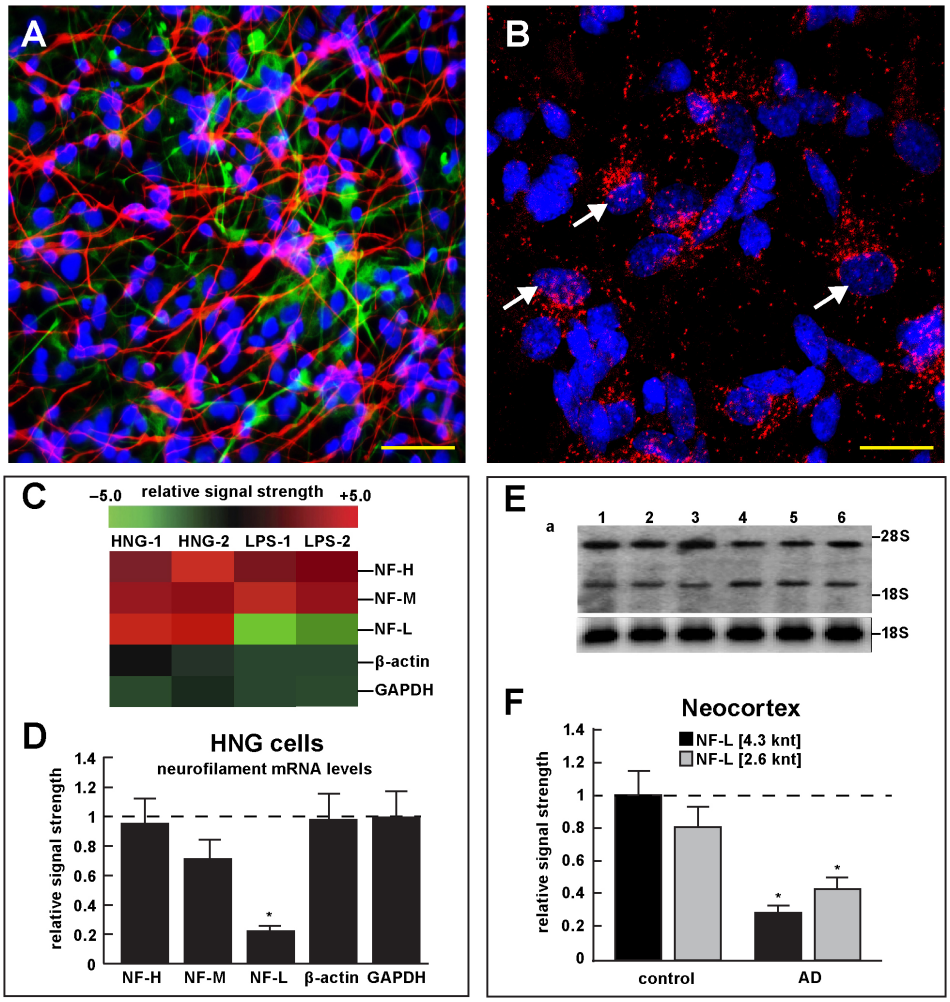

FIGURE 4 | Studies of LPS-neuronal nuclear binding in HNG cells in primary culture. **(A)** human neuronal-glial (HNG) cells in primary co-culture at 2 weeks; neurons (red stain; λ_{max} = 690 nm), DAPI (blue nuclear stain; λ_{max} = 470 nm) and GFAP (glial-specific green stain; λ_{max} = 520 nm); human neurons do not culture well in the absence of glia; neurons also show both extensive arborization and display electrical activity (unpublished; Lonza); scale bar = 20 μm; **(B)** details of association of LPS (red stain; λ_{max} = 690 nm) and nuclear DAPI (blue stain; λ_{max} = 470 nm); note affinity of red-stained LPS with blue-stained nuclei after only 48 h of co-incubation (arrows); see also **Supplementary File 1** (Details of accumulation of LPS in HNG cells in primary culture); scale bar for all photos (lower right) = 10 μm; **(C)** Neurofilament heavy, medium and light (NF-H, NF-M, and NF-L) chain abundance in control and LPS-treated HNG cells – cluster analysis of gene expression (mRNA levels); in two controls (HNG-1 and HNG-2) and in two LPS-treated samples (LPS-1, LPS-2), LPS-treated HNG cells exhibit a marked reduction in NF-L expression, a reduction that is not as apparent in NF-H or NF-M expression; NF-H, NF-M, and NF-L expression was quantified against the levels of β-actin and GAPDH in the same sample; **(D)** samples are quantified in bar graph format showing the mean and one standard deviation of all three neurofilament protein levels; there was no statistically significant change in NF-H, NF-M, β-actin, or GAPDH between control and LPS-treated HNG cells, however NF-L levels were reduced to about 0.22-fold of controls in LPS-treated HNG cells; interestingly the NF-H, NF-M, and NF-L mRNAs encode intermediate filaments of ∼60, ∼100, and ∼110 kDa, respectively, but due to extensive post-translational modifications such as phosphorylation and glycosylation, NF-H, NF-M, and NF-L exhibit higher molecular weights after SDS-PAGE (Western) analysis of ∼68, ∼160, and ∼205 kDa, respectively; a dashed horizontal line at 1.0 is included for ease of comparison; N = 3 to 5 experiments for each treatment; *$p < 0.01$ (ANOVA); **(E)** Northern blot analysis – decreased NF-L in AD – Northern analysis of total NF-L mRNA in control (lanes 1–3) and AD (lanes 4–6) temporal lobe neocortex (Brodmann A22); the position of the migration of 28S and 18S RNA (4.7 and 1.9 knt, respectively) are marked on the right of the gel (upper panel); the size of the two prominent NF-L mRNA bands detected are respectively about 4.3 and 2.6 knt in length; an 18S RNA was used as an internal control marker (lower panel); **(F)** Northern blots were quantified in bar graph format showing the mean and one standard deviation of decreased NF-L mRNA signals in AD neocortex versus age-matched controls; in AD the 2 NF-L bands [between the 28S and 18S RNA markers of part **(E)**] together are about 0.3- to 0.4-fold AD over control; *$p < 0.01$ (ANOVA).

(Jiang et al., 2017; Magalhães et al., 2017). Leakage of GI-tract neurotoxins into the systemic circulation may be a biomarker and an early indicator of progressive, age-related neuroinflammatory disorders that include AD (Lukiw, 2017; Magalhães et al., 2017; Montagne et al., 2017; Cox and Weiner, 2018; Ho et al., 2018; Sweeney et al., 2018). Indeed progressive "leakage" of LPS across the GI-tract and blood brain barrier (BBB) may be a feature accompanying normal aging (Montagne et al., 2017; Sweeney et al., 2018). Certain GI-tract microbiota may actively assist the host in moderating inflammatory neurodegeneration

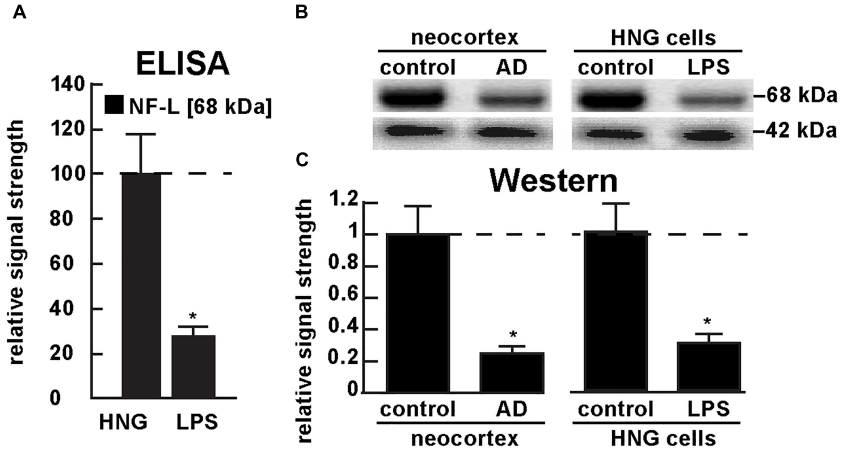

FIGURE 5 | Decreased NF-L protein in LPS-treated HNG cells and in AD: ELISA and Western analysis. **(A)** results of sandwich ELISA analysis for NF-L protein in LPS-treated HNG cells using Abbexa (abx250460; Cambridge, United Kingdom) and/or LifeSpan BioSciences (LSBio; LS-F6701; Seattle WA, United States); the 68 kDa NF-L species is a particularly abundant intermediate filament protein, however in the presence of LPS the abundance of NF-L protein was found to be reduced to about 0.3-fold of control; a dashed horizontal line at 100 is included for ease of comparison; $N = 3$ to 5 experiments per determination; $^*p < 0.01$ (ANOVA); **(B)** Western analysis of total NF-L protein (MW ∼68 kDa) in control (pool of five controls and five AD temporal lobe neocortex Brodman A22) and total NF-L protein in control and LPS-treated HNG cells (at 2 weeks of culture; see **Figure 4A**); β-actin protein (MW ∼42 kDa) was used as an internal control marker in the same sample for each determination; **(C)** Western blots were quantified in bar graph format of decreased NF-L protein abundance in AD neocortex versus age-matched controls and in LPS-treated HNG cells versus age-matched controls; a dashed horizontal line at 1.0 is included for ease of comparison; the results of decreased NF-L expression for AD over control or LPS-treated HNG cells over control are highly significant; $N = 3$ to 5 experiments; $^*p < 0.01$ (ANOVA).

by supporting the generation of short chain fatty acids (SCFAs) that can pass these barriers and subsequently interfere with the generation and aggregation of neurotoxic amyloid beta (Aβ) peptides (Cox and Weiner, 2018; Ho et al., 2018). The interactions between LPS and amyloid peptides in the brain parenchyma remain incompletely understood, especially their dynamics and potential association as they accumulate in parallel with aging both within the confines of the CNS and throughout the systemic circulation (Zhao et al., 2017a,b,c).

LPS Transit Across Biophysical and Physiological Barriers Into the CNS

Recent data regarding the contribution of neurotoxic exudates of the human GI tract microbiome to the potential initiation, development, and/or progression of AD appears to be age-related, complicated, and significant. Major bacterial species of the human GI-tract microbiome such as the Gram-negative bacillus *Bacteroides fragilis* (*B. fragilis*) secrete a unusually complex array of highly pathogenic pro-inflammatory neurotoxins which, when released from the confines of a healthy GI tract, are highly toxic to neurons of the CNS and PNS. While an environmental cause for sporadic AD has often been suggested, a strong source of powerful neurotoxins already reside within our GI tract microbiome. LPS for example represents an internally generated GI tract microbiome-derived neurotoxin capable of driving AD-type change and has enormous potential to initiate and/or propagate inflammatory neurodegeneration along the gut-brain axis. Some incompletely understood aspects of the bioavailability to the CNS of GI-tract generated neurotoxins are (i) their translocation through the GI tract and BBB that involves dynamic structures which are known to become more "leaky" with aging and disease; (ii) the direct influence of endotoxins, such as fragilysin, which targets zonula adherens protein E-cadherin and cell–cell adhesion; and (iii) the molecular exchanges between the GI tract, the systemic circulation and the BBB to access the brain parenchyma (Seong et al., 2015; Zhan et al., 2016; Montagne et al., 2017; Tsou et al., 2017; Sweeney et al., 2018). To cite recent examples from the literature: (i) BF-LPS represents an internally generated GI tract microbiome-derived neurotoxin capable of driving and emulating AD-type change *in vitro* (Zhao and Lukiw, 2018); (ii) BF-LPS has enormous potential to initiate and/or propagate inflammatory neurodegeneration along the GI tract-CNS axis (Zhao et al., 2017a); and (iii) LPS has an unusually high and remarkable affinity for the periphery of neuronal nuclei of the human neocortex (**Figures 1–4**) (Zhao et al., 2017b).

LPS and Perinuclear Association in Sporadic AD Brain

In middle-to-late-stage AD brain the perinuclear association of LPS in AD appears to be configured into "net-like" or "clathrin-like" lattice within the neuronal cell cytoplasm (**Figures 1–4**). The pyramidal neuronal nuclei of the human neocortex are among the largest known nuclei in the human CNS, often achieving diameters of 10–20 μm and occupy a large proportion of the neuronal soma, conducive to their "euchromatic" nature and extremely high rates of transcription (Colangelo et al., 2002; Clement et al., 2016). The current experimental evidence further suggests that neuronal nuclear "encasement" by LPS appears to have a deleterious effect on the free exit and/or expression

of mRNAs through the nuclear pores into the cytoplasm – previous work has shown a decreased rate of abundance of neuronal-specific DNA transcripts in sporadic AD brain (Zhao et al., 2017b). Why the NF-L mRNA exhibits the greatest down-regulated of the neurofilament triad is not known; the NF-L light chain polypeptide is the most abundant member of the neuron-specific intermediate filament family triplet, is the major component of the highly dynamic and plastic neurites, synaptic structures, and constitutes the core of the neuronal cytoskeleton. NF-L expression is also the major regulator of the caliber of the neuronal axoskeleton and essential in neuronal development, regeneration, the plasticity of the neuronal cytoskeleton, and in the creation and the maintenance of neuronal cytoarchitecture (Julien and Mushynski, 1998; Braissant, 2007; Lam et al., 2017; Goldmann, 2018; Abu-Rumeileh et al., 2018). NF-L also serves in a critical "organizer" role in axons and dendrites and contains multiple phosphorylation sites for a surprisingly large number of neuronal-enriched protein kinases, including protein kinase A, protein kinase C, cyclin-dependent kinase 5, extracellular signal regulated kinase, glycogen synthase kinase-3, and stress-activated protein kinase gamma (de Leeuw et al., 2018; Goldmann, 2018). Perturbations in NF-L phosphorylation, structure and/or function are often observed in age-related human neurodegenerative diseases including amyotrophic lateral sclerosis, Parkinson's disease and AD, and a down-regulation of NF-L mRNA and the presence of atypical insoluble twisted neurofilament deposits have long been known to be a common feature of an abnormal neurofilament network as seen in diseased brain tissues undergoing pro-inflammatory neurodegeneration (McLachlan et al., 1988; Lukiw et al., 1990; Julien and Mushynski, 1998; Abu-Rumeileh et al., 2018; Goldmann, 2018; Zhao and Lukiw, 2018). Importantly, neuron loss in AD appears not to be the reason for the observed loss in NF-L; a classic study of 22 control, non-AD dementia and AD brains indicated that the significant decrease of NF-L mRNA in AD neocortex could not be adequately accounted for by a non-specific effect of brain damage, by neuron cell loss or by neurons with neurofibrillary degeneration (McLachlan et al., 1988; Julien and Mushynski, 1998; Ginsberg et al., 2000; Zhao et al., 2017a,b).

Interestingly, there are reports that plasma levels of NF-L protein appear to be increased in AD and may be used as a reliable diagnostic marker for AD incidence and severity (Lista et al., 2017; Abu-Rumeileh et al., 2018; Hampel et al., 2018). However, the universality of this elevation and usefulness of plasma NF-L as a biomarker for AD has been recently brought into question (Lam et al., 2017; Zhou et al., 2017; Abu-Rumeileh et al., 2018). Very recently NF-L has been found to be significantly increased in the CSF of early-onset AD patients compared to younger controls; however, this change was not found in older AD groups (Lauridsen et al., 2017). These discrepancies and differential localization of neuronal- and AD-relevant molecules are reminiscent of the variation in abundance of Aβ42 peptides within AD tissues and the extracellular fluids such as the CSF that surrounds them; for example Aβ42 peptides show an inverse abundance between brain tissues and the CSF (Fagan et al., 2006; Grimmer et al., 2009). Another example is potentially pathogenic microRNAs including a pro-inflammatory miRNA-146a and several let-7 species which are differentially abundant in AD tissues versus the AD CSF (Alexandrov et al., 2012; Derkow et al., 2018). It is tempting to speculate that a differential "compartmentalization" of neuronal-associated AD-relevant molecular species may be one significant consequence of the neuropathology of the AD process.

Important Unanswered Questions

Several fundamental questions remain concerning GI-tract microbiome exudates, their compartmentalization in the GI-tract and their potential effects on the neurobiology, neuropathology, and pathogenetics of neurodegenerative and neuropsychiatric disease. For example: does the presence of LPS and envelopment of neuronal nuclei in AD significantly affect the transcription of any other brain genes or neuron-enriched transcripts besides NF-L? Why are NF-L mRNA abundances selectively affected? Perhaps because NF-L expression is from an extremely high

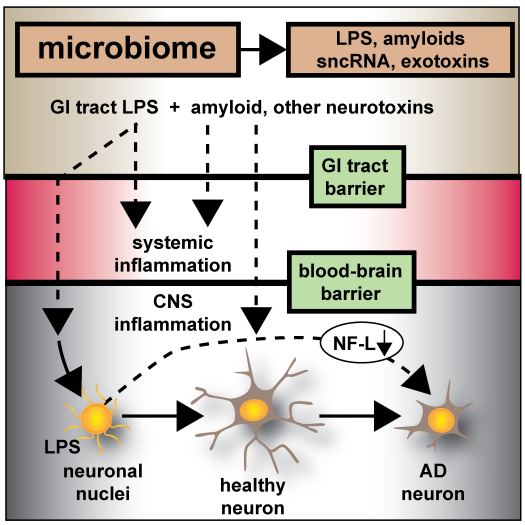

FIGURE 6 | Microbiome-derived LPS-mediated impairment of NF-L expression may contribute to atrophy of neurons and cytoskeletal disorganization that is characteristic of sporadic AD – the human GI-tract microbiome secretes a remarkably heterogeneous and complex mixture of neurotoxins including different varieties of lipooligosaccahrides (LOS), lipopolysaccharide (LPS), amyloids, small non-coding RNAs (sncRNAs) and exotoxins; recently several laboratories have provided evidence that these neurotoxins may transit GI-tract and blood-brain barriers and are present in the CNS and within aged or AD brain tissues; whether these microbiome-derived neurotoxins originate from the gastrointestinal (GI) tract microbiome, a possible brain microbiome or some dormant pathological microbiome is currently not well understood. Recent studies further suggest that the co-localization of pro-inflammatory LPS with AD-affected neuronal nuclei provides evidence that there may be a contribution of LPS to genotoxic events that support deficits in homeostatic gene expression that drive progressive AD-type change and provide support for pro-inflammatory neurodegeneration. This communication provides evidence that in both LPS-enveloped neuronal nuclei in AD neocortex and LPS-treated HNG cells in primary co-culture that there is a significant deficit in the expression of the neurofilament light-chain (NF-L), a neuron-specific cytoskeletal element known to be important in maintaining the shape and synaptic integrity of the neuron; see text and **Figures 1**–**5** for additional details.

output gene? Are the translocation of transcripts exiting the neuronal nuclei impaired by the envelopment of nuclei by LPS as they appear to be? Does a life-long exposure to certain infectious agents and their secreted neurotoxins predispose an individual to develop AD at a later age? How do the secreted neurotoxins from the human GI-tract microbiome progressively leak across biophysical and physiological barriers to access the systemic circulation and CNS compartments? Do these secreted neurotoxins interact with amyloid-beta (Aβ) peptides that increase in parallel in the aging brain? What GI-tract bacterial-derived mixtures of neurotoxins are the most pathogenic in promoting inflammatory neuro-degeneration? Can the incidence of systemic inflammation be used as a biomarker or be of prognostic value to AD and other progressive neurodegenerative diseases? Does anaerobic Gram-negative bacilli-derived LPS interact pathologically with other toxins originating from archaebacteria, fungi, protozoa, viruses, and other GI-tract resident microbes? Is there a potential synergism in their combined neurotoxic actions toward neuronal nuclei of the human CNS? Perhaps most importantly, is it possible to devise a dietary strategy that promotes the lowering of LPS secretion and optimize life-long GI-tract microbiome and CNS health to minimize the risk of developing AD as we age? Furthering our molecular and mechanistic understanding of how individual secreted components of the GI tract microbiome – affect the PNS and CNS may uncover potential and novel strategies for GI tract-based modulation of neural function and the more efficacious clinical management of terminal, age-related neurological disease.

CONCLUSION

Microbiologists, neurologists and bioinformatics researchers are still in a relatively early stage of understanding the molecular-genetic pathological signaling mechanisms that operate between the human GI-tract microbiome and the CNS of the host. Emerging evidence suggests: (i) that non-homeostatic communication along the gut-brain axis may contribute to progressive inflammatory neurodegeneration and AD-type change in the CNS; and (ii) that the GI-tract microbiome appears to have potential to contribute to neurodegenerative disease through the release, and export into the systemic circulation, of multiple, highly pro-inflammatory neurotoxic exudates predominantly from abundant species of Gram-negative anaerobic bacteria such as *Bacteroides fragilis* and other GI-tract microbes (Bhattacharjee and Lukiw, 2013; Foster et al., 2016; Zhao et al., 2017b,c; Cox and Weiner, 2018; Zhan et al., 2018; Zhao and Lukiw, 2018). The co-localization and eventual envelopment of sporadic AD-affected neuronal nuclei with a "clathrin-like" cage of GI-tract microbiome-derived LPS suggests that there exists a novel pathogenic contribution of LPS to neuron-specific gene expression and transcriptional output from AD-affected neurons (Zhao et al., 2017b). In agreement with previous reports a remarkably high proportion of the LPS signal in AD neocortex and hippocampus and in LPS-treated HNG cells are associated with neuronal nuclei (**Figures 1**–**5** and **Supplementary File 1**; Zhao et al., 2017a,b,c; Zhao and Lukiw, 2018). Recent data further indicates that patients with moderate-to-advanced sporadic AD appear to have a significantly higher population of LPS-enveloped nuclei, and a correspondingly lower amount of NF-L associated with that region of the AD brain (Zhao et al., 2017b; Zhao and Lukiw, 2018) (**Figures 5**, **6**).

In conclusion, the current work provides five novel observations: (i) that in AD neocortex LPS has a remarkable biophysical affinity for neuronal nuclei; (ii) that this action appears to selectively impair the transcriptional abundance of neuron-specific elements such as NF-L, known to be normally required for the maintenance of neuronal cytoarchitecture, synaptic connections, and the homeostatic signaling operations of neurons; (iii) that GI tract microbiome-derived neurotoxins may contribute to AD-type pathological and neuronal architectural change; (iv) that LPS-treated HNG brain cells in primary culture can recapitulate this phenomenon both at the biophysical and transcriptional level; and (v) perhaps most significantly, that LPS-treated HNG cells in primary culture can provide a highly useful experimental platform for further study on LPS effects on AD-like processes and their pathogenic consequences. These results also support the hypothesis that GI-tract derived, microbial neurotoxins such as LPS affect the efficient readout of AD-relevant neuronal-specific genetic information, such as that from the NF-L gene, and progressively contribute to cytoarchitectural aberrations, neuronal atrophy, and synaptic disorganization all of which are characteristic features of the sporadic AD brain (**Figure 6**).

ETHICS STATEMENT

Procedures involving short post-mortem interval (PMI) human tissues and human brain primary cell cultures (HNG) were followed and handled in strict accordance with the ethics review board policies at donor institutions, and the Institutional Biosafety Committee/Institutional Review Board (IBC/IRB) ethical guidelines at the LSU Health Sciences Center, LA 70112, United States (IBC No. 12323; IRB No. 6774).

AUTHOR CONTRIBUTIONS

LC, VJ, WL, and YZ preformed the experiments and analyzed the data. WL wrote the paper.

FUNDING

Research on microRNAs, pro-inflammatory, and pathogenic signaling in the Lukiw Laboratory involving the microbiome, the innate-immune response, amyloidogenesis, synaptogenesis, and neuroinflammation in AD, prion, and in other neurological diseases was supported through an unrestricted grant to the LSU Eye Center from Research to Prevent Blindness (RPB); the Louisiana Biotechnology Research Network (LBRN), and NIH grants NEI EY006311, NIA AG18031, and NIA AG038834 (WL).

ACKNOWLEDGMENTS

The experimental work in this paper was presented in part at the Vavilov Institute of General Genetics/Moscow State University Autumn 2017 Seminar Series (Институт общей генетики имени Вавилова Осень 2017 Семинар серии) in Moscow, RUSSIA October 2017 and at the Society for Neuroscience (SFN) Annual Meeting, Washington, DC, United States November 2017. We would like to express our sincere thanks to Drs L. Carver, L. Cong, J. G. Cui, F. Culicchia, C. Eicken, V. Jaber, K. Navel, A. I. Pogue, W. Poon and the late Drs. J. M. Hill and T. P. Kruck for helpful discussions in this research area, for short post-mortem interval (PMI) human brain, and HNG cells or extracts, for initial bioinformatics and data interpretation, and to D Guillot for expert technical assistance and medical artwork. Thanks are also extended to the University of California at Irvine Brain Bank, the University of Maryland Brain Bank, the Harvard Brain Tissue Resource Center, the LSU School of Medicine, and the many neuropathologists, physicians, and researchers in the United States and Canada who have provided high quality, short PMI human brain tissue fractions for scientific analysis. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging, the National Center for Research Resources, or the National Institutes of Health.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2018.00896/full#supplementary-material

REFERENCES

Abu-Rumeileh, S., Capellari, S., Stanzani-Maserati, M., Polischi, B., Martinelli, P., Caroppo, P., et al. (2018). The CSF neurofilament light signature in rapidly progressive neurodegenerative dementias. *Alzheimers Res. Ther.* 10:3. doi: 10.1186/s13195-017-0331-1

Alexandrov, P. N., Dua, P., Hill, J. M., Bhattacharjee, S., Zhao, Y., and Lukiw, W. J. (2012). microRNA (miRNA) speciation in Alzheimer's disease (AD) cerebrospinal fluid (CSF) and extracellular fluid (ECF). *Int. J. Biochem. Mol. Biol.* 3, 365–373.

Bagyinszky, E., Giau, V. V., Shim, K., Suk, K., An, S. S. A., and Kim, S. (2017). Role of inflammatory molecules in the Alzheimer's disease progression and diagnosis. *J. Neurol. Sci.* 376, 242–254. doi: 10.1016/j.jns.2017.03.031

Bhattacharjee, S., and Lukiw, W. J. (2013). Alzheimer's disease and the microbiome. *Front. Cell Neurosci.* 7:153. doi: 10.3389/fncel.2013.00153

Braissant, O. (2007). "Neurofilament proteins in brain diseases," in *New Research on Neurofilament Proteins ISBN: 1-60021-396-0*, ed. R. K. Arlen (Hauppauge, NY: Nova Science Publishers, Inc), 25–51.

Clark, I. A., and Vissel, B. (2015). Amyloid β: one of three danger-associated molecules that are secondary inducers of the proinflammatory cytokines that mediate AD. *Br. J. Pharmacol.* 172, 3714–3727. doi: 10.1111/bph.13181

Clement, C., Hill, J. M., Dua, P., Culicchia, F., and Lukiw, W. J. (2016). Analysis of RNA from Alzheimer's disease post-mortem brain tissues. *Mol. Neurobiol* 53, 1322–1328. doi: 10.1007/s12035-015-9105-6

Colangelo, V., Schurr, J., Ball, M. J., Pelaez, R. P., Bazan, N. G., and Lukiw, W. J. (2002). Gene expression profiling of 12633 genes in Alzheimer hippocampal CA1: transcription and neurotrophic factor down-regulation and up-regulation of apoptotic and pro-inflammatory signaling. *J. Neurosci. Res.* 70, 462–473. doi: 10.1002/jnr.10351

Cox, L. M., and Weiner, H. L. (2018). Microbiota signaling pathways that influence neurologic disease. *Neurotherapeutics* 15, 135–145. doi: 10.1007/s13311-017-0598-8

Cui, J. G., Li, Y. Y., Zhao, Y., Bhattacharjee, S., and Lukiw, W. J. (2010). Differential regulation of interleukin-1 receptor-associated kinase-1 (IRAK-1) and IRAK-2 by microRNA-146a and NF-κB in stressed human astroglial cells and in Alzheimer disease. *J. Biol. Chem.* 285, 38951–38960. doi: 10.1074/jbc.m110.178848

de Leeuw, R., Gruenbaum, Y., and Medalia, O. (2018). Nuclear lamins: thin filaments with major functions. *Trends Cell Biol.* 28, 34–45. doi: 10.1016/j.tcb.2017.08.004

Dendooven, T., and Luisi, B. F. (2017). RNA search engines empower the bacterial intranet. *Biochem. Soc. Trans.* 45, 987–997. doi: 10.1042/BST20160373

Derkow, K., Rössling, R., Schipke, C., Krüger, C., Bauer, J., Fähling, M., et al. (2018). Distinct expression of the neurotoxic microRNA family let-7 in the cerebrospinal fluid of patients with Alzheimer's disease. *PLoS One* 13:e0200602. doi: 10.1371/journal.pone.0200602

Fagan, A. M., Mintun, M. A., Mach, R. H., Lee, S. Y., Dence, C. S., Shah, A. R., et al. (2006). Inverse relation between in vivo amyloid imaging load and cerebrospinal fluid Aβ42 in humans. *Ann. Neurol.* 59, 512–519. doi: 10.1002/ana.20730

Foster, J. A., Lyte, M., Meyer, E., and Cryan, J. F. (2016). Gut microbiota and brain function: an evolving field in neuroscience. *Int. J. Neuropsychopharmacol.* 19:yv114. doi: 10.1093/ijnp/pyv114

Ginsberg, S. D., Hemby, S. E., Lee, V. M., Eberwine, J. H., and Trojanowski, J. Q. (2000). Expression profile of transcripts in Alzheimer's disease tangle-bearing CA1 neurons. *Ann. Neurol* 48, 77–87. doi: 10.1002/1531-8249(200007)48:1<77::AID-ANA12>3.0.CO;2-A

Goldmann, W. H. (2018). Intermediate filaments and cellular mechanics. *Cell Biol. Int.* 42, 132–138. doi: 10.1002/cbin.10879

Grimmer, T., Riemenschneider, M., Förstl, H., Henriksen, G., Klunk, W. E., Mathis, C. A., et al. (2009). Beta amyloid in Alzheimer's disease: increased deposition in brain is reflected in reduced concentration in cerebrospinal fluid. *Biol. Psychiatry* 65, 927–934. doi: 10.1016/j.biopsych.2009.01.027

Hampel, H., Toschi, N., Baldacci, F., Zetterberg, H., Blennow, K., Kilimann, I., et al. (2018). Alzheimer's disease biomarker-guided diagnostic workflow using the added value of six combined cerebrospinal fluid candidates: Aβ(1-42), total-tau, phosphorylated-tau, NFL, neurogranin, and YKL-40. *Alzheimers Dement.* 14, 492–501. doi: 10.1016/j.jalz.2017.11.015

Ho, L., Ono, K., Tsuji, M., Mazzola, P., Singh, R., Pasinetti, G. M. et al. (2018). Protective roles of intestinal microbiota derived short chain fatty acids in Alzheimer's disease-type beta-amyloid neuropathological mechanisms. *Expert Rev. Neurother.* 18, 83–90. doi: 10.1080/14737175.2018.1400909

Hofer, U. (2014). Microbiome: *B. fragilis* and the brain. *Nat. Rev. Microbiol.* 12, 76–77. doi: 10.1038/nrmicro3197

Jaber, V., Zhao, Y., and Lukiw, W. J. (2017). Alterations in micro RNA-messenger RNA (miRNA-mRNA) coupled signaling networks in sporadic Alzheimer's disease (AD) hippocampal CA1. *J. Alzheimers Dis. Parkinsonism* 7:312. doi: 10.4172/2161-0460.1000312

Jiang, C., Li, G., Huang, P., Liu, Z., and Zhao, B. (2017). The gut microbiota and Alzheimer's disease. *J. Alzheimers Dis.* 58, 1–15. doi: 10.3233/JAD-161141

Julien, J. P., and Mushynski, W. E. (1998). Neurofilaments in health and disease. *Prog. Nucleic Acid Res. Mol. Biol.* 61, 1–23. doi: 10.1016/S0079-6603(08)60823-5

Khalil, B., Morderer, D., Price, P. L., Liu, F., and Rossoll, W. (2018). mRNP assembly, axonal transport, and local translation in neurodegenerative diseases. *Brain Res.* 1693(Pt A), 75–91. doi: 10.1016/j.brainres.2018.02.018

Köhler, C. A., Maes, M., Slyepchenko, A., Berk, M., Solmi, M., Lanctôt, K. L., et al. (2016). The gut-brain axis, including the microbiome, leaky gut and bacterial translocation: mechanisms and pathophysiological role in Alzheimer's disease. *Curr. Pharm. Des.* 22, 6152–6166. doi: 10.2174/1381612822666160907093807

Lam, K. Y. C., Huang, Y., Yao, P., Wang, H., Dong, T. T. X., Zhou, Z., et al. (2017). Comparative study of different acorus species in potentiating neuronal

differentiation in cultured PC12 cells. *Phytother. Res.* 31, 1757–1764. doi: 10.1002/ptr.5904

Lauridsen, C., Sando, S. B., Møller, I., Berge, G., Pomary, P. K., Grøntvedt, G. R., et al. (2017). Cerebrospinal fluid Aβ43 is reduced in early-onset compared to late-onset Alzheimer's disease, but has similar diagnostic accuracy to Aβ42. *Front. Aging Neurosci.* 9:210. doi: 10.3389/fnagi.2017.00210

Li, Y. Y., Cui, J. G., Dua, P., Pogue, A. I., Bhattacharjee, S., and Lukiw, W. J. (2011). Differential expression of miRNA-146a-regulated inflammatory genes in human primary neural, astroglial and microglial cells. *Neurosci. Lett.* 499, 109–113. doi: 10.1016/j.neulet.2011.05.044

Lista, S., Toschi, N., Baldacci, F., Zetterberg, H., Blennow, K., Kilimann, I., et al. (2017). Diagnostic accuracy of CSF neurofilament light chain protein in the biomarker-guided classification system for Alzheimer's disease. *Neurochem. Int.* 108, 355–360. doi: 10.1016/j.neuint.2017.05.010

Lukiw, W. J. (2016a). *Bacteroides fragilis* lipopolysaccharide and inflammatory signaling in Alzheimer's disease. *Front. Microbiol.* 7:1544. doi: 10.3389/fmicb.2016.01544

Lukiw, W. J. (2016b). The microbiome, microbial-generated pro-inflammatory neurotoxins, and Alzheimer's disease. *J. Sport Health Sci.* 5, 393–396. doi: 10.1016/j.jshs.2016.08.008

Lukiw, W. J. (2017). The microbiome, microbial-generated pro-inflammatory neurotoxins, and Alzheimer's disease. *J. Sport Health Sci.* 5, 393–396. doi: 10.1016/j.jshs.2016.08.008

Lukiw, W. J., Cui, J. G., Marcheselli, V. L., Bodker, M., Botkjaer, A., Gotlinger, K., et al. (2005). A role for docosahexaenoic acid-derived neuroprotectin D1 in neural cell survival and Alzheimer disease. *J. Clin. Invest.* 115, 2774–2783. doi: 10.1172/JCI25420

Lukiw, W. J., Wong, L., and McLachlan, D. R. C. (1990). Cytoskeletal messenger RNA stability in human neocortex: studies in normal aging and in Alzheimer's disease. *Int. J. Neurosci.* 55, 81–88. doi: 10.3109/00207459008985953

Magalhães, T. N. C., Weiler, M., Teixeira, C. V. L., Hayata, T., Moraes, A. S., Boldrini, V. O., et al. (2017). Systemic inflammation and multimodal biomarkers in amnestic mild cognitive impairment and Alzheimer's disease. *Mol. Neurobiol.* 55, 5689–5697. doi: 10.1007/s12035-017-0795-9

Mancuso, C., and Santangelo, R. (2017). AD and gut microbiota modifications: the long way between preclinical studies and clinical evidence. *Pharmacol. Res.* 129, 329–336. doi: 10.1016/j.phrs.2017.12.009

McLachlan, D. R. C., Lukiw, W. J., Wong, L., Bergeron, C., and Bech-Hansen, N. T. (1988). Selective messenger RNA reduction in Alzheimer's disease. *Brain Res.* 427, 255–261. doi: 10.1016/0169-328X(88)90048-4

Minter, M. R., Taylor, J. M., and Crack, P. J. (2016). The contribution of neuroinflammation to amyloid toxicity in Alzheimer's disease. *J. Neurochem.* 136, 457–474. doi: 10.1111/jnc.13411

Montagne, A., Zhao, Z., and Zlokovic, B. V. (2017). Alzheimer's disease: a matter of blood-brain barrier dysfunction? *J. Exp. Med.* 214, 3151–3169. doi: 10.1084/jem.20171406

Seong, E., Yuan, L., and Arikkath, J. (2015). Cadherins and catenins in dendrite and synapse morphogenesis. *Cell Adh Migr.* 9, 202–213. doi: 10.4161/19336918.2014.994919

Sweeney, M. D., Sagare, A. P., and Zlokovic, B. V. (2018). Blood-brain barrier breakdown in Alzheimer disease and other neurodegenerative disorders. *Nat. Rev. Neurol.* 14, 133–150. doi: 10.1038/nrneurol.2017.188

Szablewski, L. (2018). Human gut microbiota in health and Alzheimer's disease. *J. Alzheimers Dis.* 62, 549–560. doi: 10.3233/JAD-170908

Tsou, Y. H., Zhang, X. Q., Zhu, H., Syed, S., and Xu, X. (2017). Drug delivery to the brain across the blood-brain barrier using nanomaterials. *Small* 13:1701921. doi: 10.1002/smll.201701921

Vogt, N. M., Kerby, R. L., Dill-McFarland, K. A., Harding, S. J., Merluzzi, A. P., Johnson, S. C., et al. (2017). Gut microbiome alterations in Alzheimer's disease. *Sci. Rep.* 7:13537. doi: 10.1038/s41598-017-13601-y

Yang, N. J., and Chiu, I. M. (2017). Bacterial signaling to the nervous system through toxins and metabolites. *J. Mol. Biol.* 429, 587–605. doi: 10.1016/j.jmb.2016.12.023

Zhan, X., Stamova, B., Jin, L. W., DeCarli, C., Phinney, B., and Sharp, F. R. (2016). Gram-negative bacterial molecules associate with Alzheimer disease pathology. *Neurology* 87, 2324–2332. doi: 10.1212/WNL.0000000000003391

Zhan, X., Stamova, B., and Sharp, F. R. (2018). Lipopolysaccharide associates with amyloid plaques, neurons and oligodendrocytes in Alzheimer's disease brain. *Front. Aging Neurosci.* 10:42. doi: 10.3389/fnagi.2018.00042

Zhao, Y., Bhattacharjee, S., Jones, B. M., Hill, J., Dua, P., and Lukiw, W. J. (2014). Regulation of neurotropic signaling by the inducible, NF-kB-sensitive miRNA-125b in Alzheimer's disease (AD) and in primary human neuronal-glial (HNG) cells. *Mol. Neurobiol.* 50, 97–106. doi: 10.1007/s12035-013-8595-3

Zhao, Y., Calon, F., Julien, C., Winkler, J. W., Petasis, N. A., Lukiw, W. J., et al. (2011). Docosahexaenoic acid-derived neuroprotectin D1 induces neuronal survival via secretase- and PPARγ-mediated mechanisms in Alzheimer's disease models. *PLoS One* 6:e15816. doi: 10.1371/journal.pone.0015816

Zhao, Y., Cong, L., Jaber, V., and Lukiw, W. J. (2017a). Microbiome-derived lipopolysaccharide enriched in the perinuclear region of Alzheimer's disease brain. *Front. Immunol.* 8:1064. doi: 10.3389/fimmu.2017.01064

Zhao, Y., Cong, L., and Lukiw, W. J. (2017b). Lipopolysaccharide (LPS) accumulates in neocortical neurons of Alzheimer's disease (AD) brain and impairs transcription in human neuronal-glial primary co-cultures. *Front. Aging Neurosci.* 9:407. doi: 10.3389/fnagi.2017.00407

Zhao, Y., Jaber, V., and Lukiw, W. J. (2017c). Secretory products of the human GI tract and their potential impact on Alzheimer's disease (AD): detection of lipopolysaccharide (LPS) in AD Hippocampus. *Front. Cell. Infect. Microbiol.* 7:318. doi: 10.3389/fcimb.2017.00318

Zhao, Y., and Lukiw, W. J. (2015). Microbiome-generated amyloid and potential impact on amyloidogenesis in Alzheimer's disease (AD). *J. Nat. Sci.* 1:e138.

Zhao, Y., and Lukiw, W. J. (2018). Microbiome-mediated upregulation of microRNA-146a in sporadic Alzheimer's disease. *Front. Neurol.* 9:145. doi: 10.3389/fneur.2018.00145

Zhou, W., Zhang, J., Ye, F., Xu, G., Su, H., Su, Y., et al. (2017). Alzheimer's disease neuroimaging initiative. plasma neurofilament light chain levels in Alzheimer's disease. *Neurosci. Lett.* 650, 60–64. doi: 10.1016/j.neulet.2017.04.027

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 13 March 2019
doi: 10.3389/fnins.2019.00220

Genetic Overlap Between Alzheimer's Disease and Bipolar Disorder Implicates the MARK2 and VAC14 Genes

Ole Kristian Drange[1,2], Olav Bjerkehagen Smeland[3,4], Alexey A. Shadrin[3], Per Ivar Finseth[5], Aree Witoelar[3,4], Oleksandr Frei[3], Psychiatric Genomics Consortium Bipolar Disorder Working Group, Yunpeng Wang[3,4], Sahar Hassani[3,6], Srdjan Djurovic[6,7], Anders M. Dale[8,9,10] and Ole A. Andreassen[3,4]*

[1] Department of Research and Development, Department of Mental Health, Norwegian University of Science and Technology, Trondheim, Norway, [2] Department of Østmarka, Division of Mental Health Care, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway, [3] Norwegian Centre for Mental Disorders Research, KG Jebsen Centre for Psychosis Research, Institute of Clinical Medicine, University of Oslo, Oslo, Norway, [4] Norwegian Centre for Mental Disorders Research, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway, [5] Department of Brøset, Division of Mental Health Care, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway, [6] Department of Medical Genetics, Oslo University Hospital, Oslo, Norway, [7] Norwegian Centre for Mental Disorders Research, KG Jebsen Centre for Psychosis Research, Department of Clinical Science, University of Bergen, Bergen, Norway, [8] Center for Multimodal Imaging and Genetics, Department of Radiology, University of California, San Diego, La Jolla, CA, United States, [9] Department of Neurosciences, University of California, San Diego, La Jolla, CA, United States, [10] Department of Psychiatry, University of California, San Diego, La Jolla, CA, United States

OPEN ACCESS

Edited by:
Efthimios M. C. Skoulakis, Alexander Fleming Biomedical Sciences Research Center, Greece

Reviewed by:
Petr A. Slominsky, Institute of Molecular Genetics (RAS), Russia
Ioannis Sotiropoulos, University of Minho, Portugal

***Correspondence:**
Ole Kristian Drange
ole.kristian.drange@gmail.com

Specialty section:
This article was submitted to Neurodegeneration, a section of the journal Frontiers in Neuroscience

Received: *30 January 2018*
Accepted: *26 February 2019*
Published: *13 March 2019*

Citation:
Drange OK, Smeland OB, Shadrin AA, Finseth PI, Witoelar A, Frei O, Psychiatric Genomics Consortium Bipolar Disorder Working Group, Wang Y, Hassani S, Djurovic S, Dale AM and Andreassen OA (2019) Genetic Overlap Between Alzheimer's Disease and Bipolar Disorder Implicates the MARK2 and VAC14 Genes. Front. Neurosci. 13:220. doi: 10.3389/fnins.2019.00220

Background: Alzheimer's disease (AD) and bipolar disorder (BIP) are complex traits influenced by numerous common genetic variants, most of which remain to be detected. Clinical and epidemiological evidence suggest that AD and BIP are related. However, it is not established if this relation is of genetic origin. Here, we applied statistical methods based on the conditional false discovery rate (FDR) framework to detect genetic overlap between AD and BIP and utilized this overlap to increase the power to identify common genetic variants associated with either or both traits.

Methods: We obtained genome wide association studies data from the International Genomics of Alzheimer's Project part 1 (17,008 AD cases and 37,154 controls) and the Psychiatric Genetic Consortium Bipolar Disorder Working Group (20,352 BIP cases and 31,358 controls). We used conditional QQ-plots to assess overlap in common genetic variants between AD and BIP. We exploited the genetic overlap to re-rank test-statistics for AD and BIP and improve detection of genetic variants using the conditional FDR framework.

Results: Conditional QQ-plots demonstrated a polygenic overlap between AD and BIP. Using conditional FDR, we identified one novel genomic locus associated with AD, and nine novel loci associated with BIP. Further, we identified two novel loci jointly associated with AD and BIP implicating the *MARK2* gene (lead SNP rs10792421, conjunctional FDR = 0.030, same direction of effect) and the *VAC14* gene (lead SNP rs11649476, conjunctional FDR = 0.022, opposite direction of effect).

Conclusion: We found polygenic overlap between AD and BIP and identified novel loci for each trait and two jointly associated loci. Further studies should examine if the shared loci implicating the *MARK2* and *VAC14* genes could explain parts of the shared and distinct features of AD and BIP.

Keywords: Alzheimer's disease, bipolar disorder, GWAS, pleiotropy, cognitive symptoms, affective symptoms, MARK2, VAC14

INTRODUCTION

About a century ago, Alois Alzheimer and Emil Kraepelin described the historical equivalents of AD and BIP (Alzheimer, 1907; Kraepelin, 1921). Still their etiologies are incompletely understood and no curative treatments exist (Grande et al., 2016; Scheltens et al., 2016). The Global Burden of Disease study ranks AD and BIP among the top thirty causes of years lived with disability worldwide (Vos et al., 2016).

Alzheimer's disease is a neurodegenerative disorder (Jack et al., 2013) usually presenting in late adult life (Koedam et al., 2010), while BIP is considered a neurodevelopmental disorder (Sanches et al., 2008; O'Shea and McInnis, 2016) with average age at onset in early adult life (Baldessarini et al., 2010). Yet, epidemiological, pathophysiological, and clinical data suggest that AD and BIP could be related. A recent meta-analysis reports an odds ratio of 2.4 (95% CI 1.4–4.1) for dementia of all causes among patients with BIP (Diniz et al., 2017). The risk of dementia is higher among patients with BIP compared to patients with arthritis, diabetes, and schizophrenia (Kessing et al., 1999; Kessing and Nilsson, 2003). Among patients with BIP, treatment with lithium is associated with a reduced risk of dementia (Kessing et al., 2010; Gerhard et al., 2015) and AD (Nunes et al., 2007) in most, but not all (Cheng et al., 2017), observational studies. Among patients with AD or mild cognitive impairment, a meta-analysis of randomized controlled studies found that lithium decreased cognitive decline (Matsunaga et al., 2015). Shared pathophysiological processes between AD and BIP are reported in the kynurenine pathway (Miller et al., 2006; Myint et al., 2007; Rahman et al., 2009; Gulaj et al., 2010; Maddison and Giorgini, 2015; Savitz et al., 2015). There is also evidence of inflammatory processes in both conditions (Goldstein et al., 2009; Antonio et al., 2015; Heneka et al., 2015). Further, euthymic patients with BIP have impairments of episodic memory (Torres et al., 2007) and executive dysfunction (Torres et al., 2007; Martino et al., 2015), which are also core symptoms of AD (Gold and Budson, 2008; Godefroy et al., 2016).

Despite several lines of evidence suggesting a relation between AD and BIP, it is not established if the conditions have a shared genetic basis. AD and BIP are in most cases complex traits, i.e., they are influenced by several genetic and environmental factors. Twin studies estimate the heritability of AD and BIP to 60% or higher (McGuffin et al., 2003; Kieseppä and Partonen, 2004; Gatz et al., 2006; Lichtenstein et al., 2009). Genome wide association studies (GWASs) are the gold standard for hypothesis-free assessment of associations between complex traits and common genetic variants (Corvin et al., 2010). The common variants refer to single nucleotide polymorphisms (SNPs) with minor allele frequencies > 1–5%. The power of a GWAS is a function of study sample size and the genetic architecture of the trait (i.e., the narrow-sense heritability, the number of causal variants, their effect sizes, and population frequencies) (Schork et al., 2016; Frei et al., 2018). AD and BIP are considered highly polygenic (Purcell et al., 2009; Escott-Price et al., 2015), and ∼1/3 of their heritability can be explained by SNPs with tiny effect sizes that are not individually detectable given the power of current GWASs (Lee et al., 2011, 2013; Ridge et al., 2013, 2016).

With the current sample sizes, however, the power of GWASs can be boosted by leveraging polygenic overlap between complex traits (Andreassen et al., 2013a,b, 2015). Shared genetic influences are common among complex traits (Visscher et al., 2017). Statistical methods based on the conditional FDR framework can detect polygenic overlap between complex traits and utilize this polygenic overlap to increase the power to identify common genetic variants associated with each trait and jointly with two or more traits (Andreassen et al., 2013a,b, 2015). We aimed to use these methods to identify the shared genetic basis between AD and BIP.

MATERIALS AND METHODS

Data Sources

We obtained summary statistics (i.e., effect sizes and corresponding *p*-values for all SNPs) from the IGAP (Lambert et al., 2013) and the PGC2-BIP (Stahl et al., 2019).

International Genomics of Alzheimer's Project

The IGAP is a two-stage study. We used data from stage 1 of the study, which is based upon four previously published GWASs [The European Alzheimer's Disease Initiative (Dreses-Werringloer et al., 2008; Heath et al., 2008), the Alzheimer Disease Genetics Consortium (Jun et al., 2010), The Cohorts for Heart and Aging Research in Genomic Epidemiology consortium (Psaty et al., 2009), The Genetic and Environmental Risk in AD consortium (Harold et al., 2009)] on 17,008 AD cases and 37,154 controls of European ancestry. The IMPUTE2 (Howie et al., 2009) or MaCH/Minimac (Li et al., 2010) software were used to impute SNPs from the European ancestry haplotypes

Abbreviations: AD, Alzheimer's disease; BIP, bipolar disorder; FDR, false discovery rate; GWAS, genome wide association study; IGAP, International Genomics of Alzheimer's Project; LD, linkage disequilibrium; LDSR, Linkage disequilibrium score regression; PGC2-BIP, Psychiatric Genetic Consortium 2 Bipolar Disorder Working Group; SNP single nucleotide polymorphism; QQ, quantile-quantile.

in the 1000 Genome Project (Altshuler et al., 2010). In stage 2 of the study, SNPs with p-values $< 1 \times 10^{-3}$ from stage 1 were selected for genotyping in independent samples. We did not use data from stage 2 of the study since the conditional FDR method require genome-wide summary statistics which are not inflated.

Diagnoses of AD in the sub-studies of IGAP were in most cases made clinically according to the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's disease and Related Disorders Association criteria (McKhann et al., 1984) or the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1994) criteria, or post mortem according to the National Institute of Ageing-Regan criteria (Newell et al., 1999).

Informed consents were provided from all participants, or, in the case of substantial cognitive impairment, from caregivers, legal guardians, or other proxies. The sub-studies were approved by local ethic committees.

For further details, we refer to the original publication (Lambert et al., 2013).

Psychiatric Genetic Consortium 2 Bipolar Disorder Working Group

The PGC2-BIP is a GWAS based upon 32 sub-studies on 20,352 BIP cases and 31,358 controls of European ancestry. Arrays for genotyping were chosen by each sub-study. The *Ricopoli* pipeline[1] was used to standardize quality control, imputation, and analyses of genotypic data from all samples except one. SNPs were excluded by the following criteria: missingness in > 5 (before sample removal) or 2% (after sample removal), p-value for Hardy–Weinberg equilibrium $< 1 \times 10^{-10}$ in cases or $<1 \times 10^{-6}$ in controls, missingness difference between cases and controls $> 2\%$, or autosomal heterozygosity deviation ($|F_{het}| > 0.2$). Individuals with $> 2\%$ missing genotypes were also excluded. The IMPUTE2 (Howie et al., 2009) and SHAPEIT2 (Delaneau et al., 2012) software were used for imputation.

Diagnoses of BIP were established by clinical interviews or obtained from hospital record data according to the Diagnostic and Statistical Manual of Mental Disorders 4th edition (American Psychiatric Association, 1994), the International Classification of Diseases 9th revision (World Health Organization, 1977), or the International Classification of Diseases 10th revision (World Health Organization, 1992).

Informed consents were provided from all participants. The sub-studies were approved by local ethical committees.

For further details, we refer to the original publication (Stahl et al., 2019).

Data Availability

Data from the IGAP[2] and PGC2-BIP[3] studies are publicly available for download.

[1]https://github.com/Nealelab/ricopili/wiki

[2]http://web.pasteur-lille.fr/en/recherche/u744/igap/igap_download.php

[3]https://www.med.unc.edu/pgc/results-and-downloads

Statistical Analyses

Conditional Quantile–Quantile (QQ)-Plots

We used conditional QQ-plots to visually assess pleiotropic enrichment. A conditional QQ-plot displays the distribution of p-values for the first trait, e.g., AD, conditioned on association levels for the second trait, e.g., BIP. Pleiotropic enrichment is present if the degree of leftward shift from the expected null line for the first trait is dependent on the degree of association with the second trait. For further details, we refer to previous studies (Andreassen et al., 2013a,b, 2015) and **Supplementary Methods 1.1**.

Conditional False Discovery Rate (condFDR)

The enrichment observed in conditional QQ-plots can be translated to FDR for each SNP. We used the conditional false discovery rate (condFDR) to improve power to detect SNPs associated with AD given associations with BIP and *vice versa*. condFDR is defined as "the posterior probability that a given SNP is null for the first trait given that the p-values for both traits are as small or smaller than the observed p-values" (Andreassen et al., 2015). We denoted condFDR for AD given associations with BIP as $\text{condFDR}_{(AD|BIP)}$ and for BIP given association with AD as $\text{condFDR}_{(BIP|\ AD)}$ and considered values < 0.01 significant. For further details, we refer to previous studies (Andreassen et al., 2013a,b, 2015) and **Supplementary Methods 1.2**.

Conjunctional False Discovery Rate (conjFDR)

We used conjunctional FDR (conjFDR) to identify SNPs jointly associated with AD and BIP. conjFDR is defined as "the posterior probability that a SNP is null for either phenotype or both simultaneously, given the p-values for both traits are as small or smaller than the observed p-values" (Andreassen et al., 2015). After repeating the condFDR procedure for both traits, we identified shared loci at conjFDR < 0.05, which is given by the maximum between the condFDRs for both traits. Hence, the conjFDR analysis is a conservative approach requiring that loci exceed a condFDR significance threshold for two traits simultaneously. For further details, we refer to previous studies (Andreassen et al., 2013a,b, 2015) and **Supplementary Methods 1.3**.

Conditional and Conjunctional Manhattan Plots

We constructed conditional Manhattan plots to visualize the chromosomal location of SNPs with $\text{condFDR}_{(AD|BIP)}$ (**Supplementary Figure 1**) and $\text{condFDR}_{(BIP|AD)} < 0.01$ (**Supplementary Figure 2**). We constructed a similar plot for SNPs jointly associated with AD and BIP at a conjFDR < 0.05 (**Figure 2**).

Assessment of Novelty

To determine if a locus was novel, we first checked that the p-value(s) for the implicating variant was $> 5 \times 10^{-8}$ in the original GWAS(s). Further, we used LDlink (Machiela and Chanock, 2015) to exclude variants which are in LD ($r^2 > 0.1$) with any of the genome-wide significant hits in the original GWAS(s). Finally, we conducted a search on PubMed using the term ("SNP id" OR "gene name") AND ("Bipolar

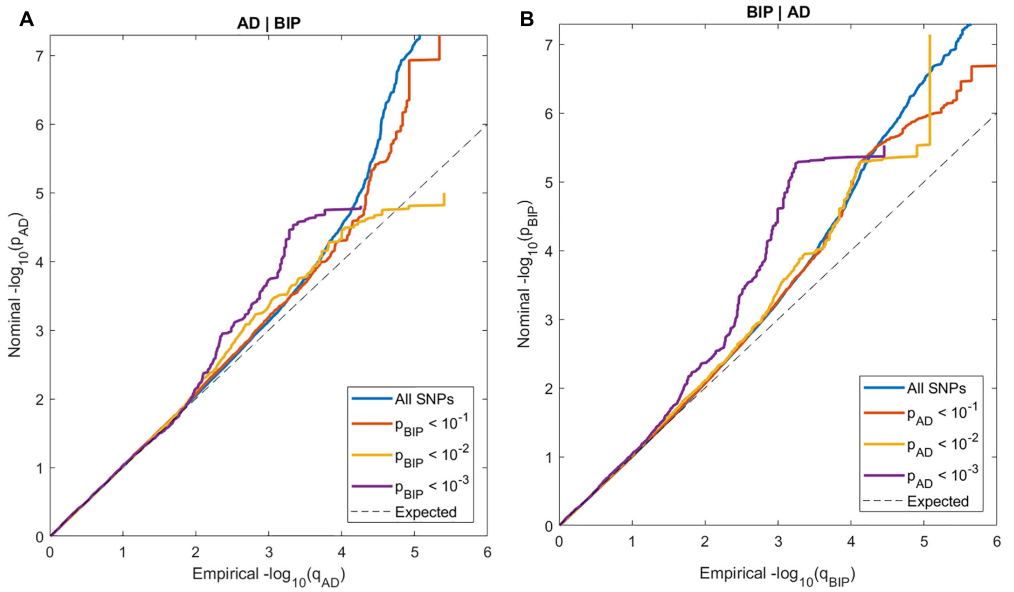

FIGURE 1 | Conditional QQ-plots of nominal p-values at y-axis and 1 - empirical cumulative distribution function on *x*-axis for **(A)** Alzheimer's disease (AD) with lines representing strata of SNPs according to their degree of association with bipolar disorder (BIP) and **(B)** BIP with lines representing strata of SNPs according to their degree of association with AD.

Disorder"[Mesh] OR "Alzheimer Disease"[Mesh]) to exclude that the variants or implicated genes have been associated with AD or BIP at genome-wide significance in previous GWASs.

Cerebral Gene Expression Across Lifespan of the Implicated Loci

The Human Brain Transcriptome (HBT) project[4] used postmortem brain tissue from over 1,340 samples to provide genome-wide exon-level transcriptome data in 16 cerebral regions (Kang et al., 2011). We obtained figures from the HBT project on gene expression in different cerebral areas as a function of age (i.e., from embryonic life through late adulthood) for the nearest genes to the loci jointly associated with AD and BIP.

Control of Spurious Enrichment

We randomly chose one SNP in each LD block ($r^2 > 0.1$), and calculated the average empirical cumulative distribution function (ecdf) by using the *p*-values obtained through 200 iterations. SNPs within the major histocompatibility complex region (defined as chr6:25652429–33368333) and the apolipoprotein E (*APOE*) gene (chr19:44909039–45912650), and SNPs in LD ($r^2 > 0.1$) with these SNPs, were excluded from the analyses due to their complex LD structure (de Bakker and Raychaudhuri, 2012) and known association to AD (Lambert et al., 2010; Scheltens et al., 2016), which could bias the estimates of enrichment. Further, we used LD-independent ($r^2 < 0.1$) intergenic SNPs, which are depleted of true associations, to calculate an inflation factor value (Wang et al., 2016a). We divided all test statistics on this value to control for genomic inflation.

Cross-Trait Linkage Disequilibrium Score Regression (LDSR)

We calculated the degree of genetic correlation between AD and BIP using cross-trait LD score regression (LDSR) (Bulik-Sullivan et al., 2015). For details, we refer to **Supplementary Materials 1.4**.

Ethics Statement

All GWASs performed and investigated in the present study were approved by the local ethics committees, and informed consent was obtained from all participants. Furthermore, the Norwegian Institutional Review Board for the South-East Norway Region has evaluated the methods used in the current study and found that no additional institutional review board approval was needed because no individual data were used (ref. 2011/1980).

RESULTS

Pleiotropic Enrichment

In the conditional QQ-plots, we observed enrichment of associations with AD given increasing SNP associations with BIP, and *vice versa* (**Figure 1**). These findings indicate polygenic overlap between AD and BIP across common genetic variants.

Improving Genetic Discovery Using Conditional FDR

We then leveraged the pleiotropic enrichment observed in conditional QQ-plots to boost SNP discovery in both traits using condFDR.

[4]http://hbatlas.org

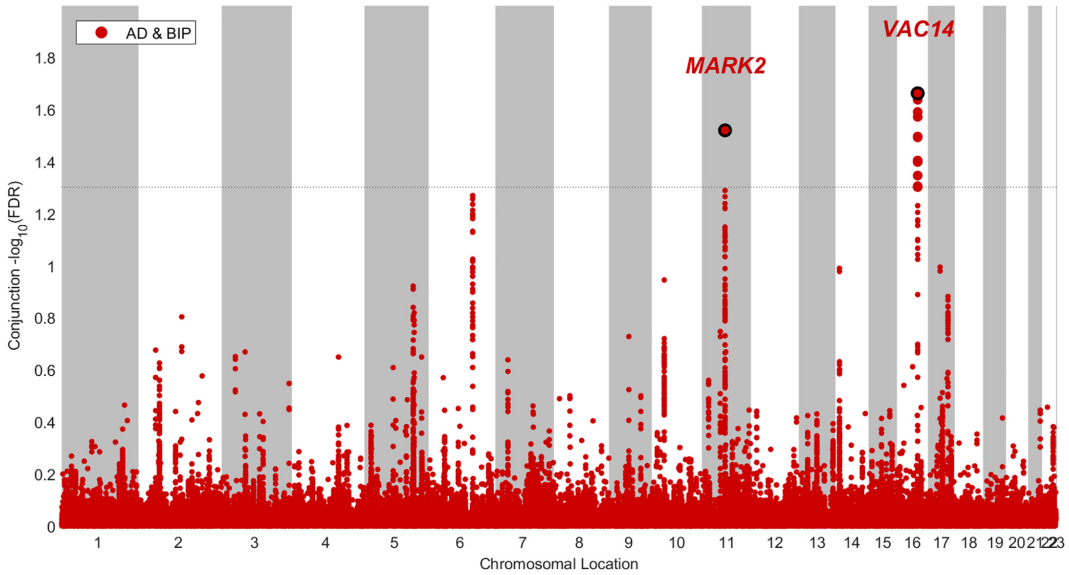

FIGURE 2 | Conjunctional Manhatton plot of loci jointly associated with Alzheimer's disease (AD) and bipolar disorder (BIP) at a conjuntional false discovery rate < 0.05.

We identified 22 SNPs clumped into 19 independent loci at condFDR$_{(AD|BIP)}$ < 0.01 (**Supplementary Table 1**). The chromosomal locations of the nearest genes are visualized in a conditional Manhattan plot (**Supplementary Figure 1**). Red annotations represent the four loci with a lower conditional than unconditional FDR. Of these four loci, two loci have uncorrected p-values $> 5 \times 10^{-8}$ in the original GWAS and are thus not identified by traditional methods; *NDUFS3* (rs71475924, intron variant) and *MTSS1L* (rs12597717, intron variant). The signal in *NDUFS3* was driven by one single SNP and is thus probably a spurious association.

Further, we identified 24 SNPs within 24 loci at a condFDR$_{(BIP|AD)}$ < 0.01 (**Supplementary Table 2**). As visualized in the conditional Manhatton plot (**Supplementary Figure 2**), 17 loci had a lower conditional than unconditional FDR. Of these 17 loci, 10 variants have uncorrected p-values $> 5 \times 10^{-8}$ in the original GWAS and are thus not identified by traditional methods; *LOC105378763* (rs1889778, intron variant), *CNTNAP5* (rs13011184, intron variant), *KIAA1109* (rs45605540, intron variant), *SSBP2* (rs7707981, intron variant), *AK091365* (rs2388334, no genic locational annotation), *RCOR2* (rs4980532, intron variant), *STARD9* (rs4447398, intron variant), *GRIN2A* (rs11647445, intron variant), *THRA* (rs61554907, intron variant), and *PRKCA* (rs7406066, intron variant). However; the *CNTNAP5* gene has previously been associated with the posterior cortical atrophy variant of AD at genome-wide significance (Schott et al., 2016) and with BIP (Djurovic et al., 2010).

Identification of Shared Loci

Finally, we applied conjFDR to assess for SNPs jointly associated with AD and BIP. We used effect sizes from the original data sources to determine the allelic direction of effects in both traits.

We identified two SNPs at two loci at a conjFDR$_{(AD\&BIP)}$ < 0.05 (**Table 1** and **Figure 2**). A 2 kb upstream variant at *MARK2* (rs10792421) was associated with AD and BIP with the same direction of effect on AD and BIP [conjFDR$_{(AD\&BIP)}$ = 0.030, z-score$_{(AD)}$ = 3.99, z-score$_{(BIP)}$ = 4.74]. *MARK2* is widely expressed in the developing and adult human brain (**Supplementary Figure 3**). An intronic variant within *VAC14* (rs11649476) was associated with AD and BIP with opposite directions of effect in AD and BIP [conjFDR$_{(AD\&BIP)}$ = 0.022, z-score$_{(AD)}$ = −4.35, z-score$_{(BIP)}$ = 4.18]. *VAC14* is also widely expressed in the developing and adult human brain (**Supplementary Figure 4**). Both SNPs have p-values $> 5 \times 10^{-8}$ for both traits in the original GWASs and are thus not identified by traditional methods.

Genetic Correlation

We estimated that there is no overall genetic correlation between AD and BIP according to LDSR (r_g = −0.0222, SE = 0.0519, p = 0.669).

DISCUSSION

We used statistical methods based on the condFDR framework and showed that AD and BIP have a shared genetic basis. Our study adds new insights into the relation between AD and BIP by finding polygenic overlap, one novel locus associated with AD and nine novel loci associated with BIP when conditioned on associations with the other trait, and two novel loci jointly associated with both traits.

A polygenic overlap between AD and BIP could implicate shared genetic influences as a part of the explanation to the epidemiological (Diniz et al., 2017), pathophysiological

TABLE 1 | SNPs with related genes jointly associated with Alzheimer's disease (AD) and bipolar disorder (BIP) at a conjunctional false discovery rate (conjFDR$_{(AD\&BIP)}$) < 0.05.

SNP	Chr. region	Position	Closest gene	Location relative to the closest gene	*P*-value$_{(AD)}$	*P*-value$_{(BIP)}$	conjFDR$_{(AD\&BIP)}$	Effective/ other allele	Direction of effect in AD/BIP
rs10792421	11q13.1	63605177	MARK2	Upstream	6.68E-05	2.16E-05	3.02E-02	G/A	+/+
rs11649476	16q22.2	70736752	VAC14	Intronic	1.35E-05	2.98E-05	2.18E-02	T/C	−/+

(Goldstein et al., 2009; Heneka et al., 2015), and clinical (Gold and Budson, 2008; Martino et al., 2015; Godefroy et al., 2016) links between the diseases. However, we do not find an overall genetic correlation as assessed with cross-trait LDSR (Bulik-Sullivan et al., 2015). Also, one of the two jointly associated SNPs demonstrates effects in opposite directions. These findings are compatible with a scenario where the polygenic overlap between AD and BIP involves a mixed direction of effects of the implicated SNPs yielding no genome-wide correlation (Frei et al., 2018). Thus, absence of an overall genetic correlation between brain disorders, as evident for several traits (including AD and BIP) in the study of Anttila et al. (2018), does not imply lack of genetic overlap.

The loci implicating the *MARK2* and *VAC14* genes were jointly associated with AD and BIP (**Table 1**). Both genes are widely expressed in the human brain throughout life (**Supplementary Figures 3, 4**), which implies a spatial and temporal relation to both neurodevelopmental and neurodegenerative processes. The locus implicating the *MARK2* gene (rs10792421) had a concordant direction of effect in both traits (**Table 1**). The *MARK2* gene encodes the microtubule affinity regulating kinase 2 (MARK2). The kinase is involved in a diversity of neuronal cellular processes, including neuronal migration, and tau phosphorylation (Matenia and Mandelkow, 2009). Migration of immature neurons is necessary for corticogenesis (Kon et al., 2017). BIP is considered a neurodevelopmental disorder partly because of previous findings of cortical cell migration abnormalities (Sanches et al., 2008; O'Shea and McInnis, 2016). Abnormal neuronal migration might also be involved in later stages of life among patients with AD (Reiner et al., 2009). Tauopathy is one of the pathophysiological hallmarks of AD (Jack et al., 2013). Gu G.J. et al. (2013) demonstrated that MARK2 increases the phosphorylation of tau *in situ* and found interactions between MARK2 and tau in postmortem human AD brain tissue. The role of tauopathy has also been explored in BIP. A study of cerebrospinal fluid among younger patients with BIP (Jakobsson et al., 2013) and a similar study of elderly patients with BIP and mild cognitive impairment (Forlenza et al., 2016) did not find any evidence of tauopathy. However, in another study, the total to phosphorylated tau ratio was reduced among patients with BIP carrying the risk allele of a common variant related to the previously discovered BIP risk gene *CACNA1C* (Jakobsson et al., 2016). A similar reduction was not found among healthy controls carrying the same risk allele. These findings suggest an alteration in the regulation of tau phosphorylation in carriers of the risk allele that is restricted to patients with BIP. Further studies should explore whether interactions with other genes involved in regulation of tau phosphorylation, like the *MARK2* gene, could explain the specificity of the finding to patients with BIP. Lithium has several molecular targets including inhibition of glycogen synthase kinase 3β (Freland and Beaulieu, 2012). Evidence is conflicting on whether glycogen synthase kinase 3β in turn inhibits or activates MARK2 (Kosuga et al., 2005; Timm et al., 2008). Consequently, it is unknown whether treatment with lithium could result in reduced or increased phosphorylation of tau among carriers of the common variant related to the *MARK2* gene.

The intronic variant within *VAC14* (rs11649476) was related to AD and BIP with opposite directions of effects. The same variant was shared between BIP and intelligence with concordant direction of effects in a recent study using conjunctional FDR (Smeland et al., 2019). *VAC14* encodes a part of the PIKfyve protein kinase complex, which phosphorylates phosphatidylinositol 3-phosphate [PI(3)P] to phosphatidylinositol 3,5-bisphosphate [PI(3,5)P_2] (McCartney et al., 2014). PI(3,5)P_2 is involved in endosomal homeostasis (Di Paolo and De Camilli, 2006). A null mutation of *VAC14* in a mouse model resulted in perinatal death and massive neurodegeneration with vacuolated neurons (Zhang et al., 2007). Amyloid precursor protein (APP) is a transmembrane protein involved in the pathophysiology of AD (O'Brien and Wong, 2011). Balklava et al. (2015) found that APP interacts with the PIKfyve complex to maintain endosomal homeostasis in *C. elegans*. They postulated that aberrant processing of APP contributes to the pathophysiology of AD through a cascade of reduced activation of PIKfyve, reduced levels of PI(3,5)P_2, endosomal dysfunction, and reduced clearance of beta amyloid. Another example of the relationship between the processing of phosphoinositides and APP comes from a study of Miranda et al. (2018). They found that inhibition of Vps34, a kinase phosphorylating phosphatidylinositol (PI) to PI(3)P, causes endolysosomal dysfunction with secretion of exosomes containing APP C-terminal fragments. Knowles et al. (2017) recently reported that serum levels of PI, the precursor of phosphoinositides like PI(3)P and PI(3,5)P_2, is negatively associated with a proxy of genetic susceptibility to BIP.

Some of the genes implicated by the novel loci identified by conditional FDR analyses (**Supplementary Tables 1, 2** and **Supplementary Figures 1, 2**) also relate to known pathophysiological and clinical features of AD and BIP. The *PRKCA* gene encodes the protein kinase C alpha (PKCa). PKCa is described in amyloid plaque of patients with AD (Clark et al., 1991) where it could contribute to reduced synaptic activity (Alfonso et al., 2016). The *PRKCA* gene is higher

expressed in bipolar mania compared to unipolar depression (Wang et al., 1999), and is lower expressed in fibroblasts of patients with BIP treated with lithium compared to those treated with other medications (Kittel-Schneider et al., 2016). Common genetic variants implicating the *PRKCA* gene are in healthy individuals associated to impairment of episodic memory (MacLeod and Donaldson, 2014). Variants within the *KIAA1109* gene are in family studies associated with multi-system syndromes characterized by impaired neurodevelopment (Alazami et al., 2015; Gueneau et al., 2018), while the *MTSS1L* gene is associated with neurodegeneration in a consanguineous family study (Alazami et al., 2015). The *STARD9* gene is necessary for spindle assembly during cell division in human development, and a mutation in the gene might cause a syndrome with intellectual disability (Okamoto et al., 2017). The locus implicating the *AK091365* gene was previously associated with general cognitive function when conditioned on association with schizophrenia (Smeland et al., 2017), which in turn has a high genetic correlation with BIP (Bulik-Sullivan et al., 2015). The *SSBP2* gene encodes the single strand DNA binding protein 2, which protects telomeres in a mouse model (Gu p. et al., 2013). In a Mendelian randomization study, Zhan et al. (2015) found that telomere length is causally related to AD. Telomere length is probably not reduced in most patients with BIP (Colpo et al., 2015; Darrow et al., 2016), however; one study found that patients with BIP treated with lithium had longer telomeres compared to patients not receiving lithium (Powell et al., 2017). The *RCOR2* gene product is related to cortical development (Wang et al., 2016b) and inflammation (Alvarez-López et al., 2014) in mice. The *GRIN2A* gene encodes the GluN2A subunit of the *N*-methyl-D-aspartate (NMDA) receptor. The NMDA receptor is central for synaptic plasticity and learning (Li and Tsien, 2009). Memantine, an NMDA receptor antagonist, probably reduces cognitive decline (Reisberg et al., 2003; Howard et al., 2012) and neuropsychiatric symptoms (Maidment et al., 2008) in AD. Ketamine, another NMDA receptor antagonist, can give short term remission of depression in BIP when used as an add-on to mood stabilizers (Diazgranados et al., 2010; Zarate et al., 2012). Mutations in *GRIN2A* are previously associated with a range of neuropsychiatric phenotypes including mental retardation, epilepsy, schizophrenia, and BIP (Itokawa et al., 2003; Yuan et al., 2015).

Some of the genes implicated both at genome-wide significance in previously GWASs and by conditional FDR in the present study also have pathophysiological and clinical plausibility. The expression of *TRANK1* is decreased in induced pluripotent stem cells derived neurons carrying the common variant found in our study (rs9834970). Decreased expression of *TRANK1* alters the expression of other genes related to neuronal development and differentiation (Jiang et al., 2018). Chronic treatment with sodium valproate, a mood stabilizer used in BIP (Macritchie et al., 2001), normalizes the expression of *TRANK1* (Jiang et al., 2018). The *CNTNAP5* gene encodes a transmembrane protein of the neurexin family, which is related to cellular adhesion and intercellular communication (Traut et al., 2006). Common variants implicating *CNTNAP5* have previously been associated with the posterior cortical atrophy variant of AD (Schott et al., 2016), BIP (Djurovic et al., 2010), and response to antipsychotic treatment in schizophrenia (Yu et al., 2018), while rare variants within *CNTNAP5* have previously been associated with autism spectrum disorders (Pagnamenta et al., 2010). The *NCAN* gene is involved in neuronal adhesion and migration (Raum et al., 2015). Common variants implicating *NCAN* are associated with cognitive performance (Raum et al., 2015) and limbic gray matter volumes (Dannlowski et al., 2015) in healthy individuals, while a rare variant is associated with dyslexia (Einarsdottir et al., 2017).

Further experimental studies should examine the implications of our findings. It is unknown if the loci implicated by condFDR and conjFDR relate to altered levels of gene expression, pathophysiological processes (e.g., impaired neuronal migration, tauopathy, and disturbed endosomal homeostasis), clinical features (e.g., cognitive and affective symptoms), and treatment response to lithium among patients with AD and BIP. Further, it is unknown if the loci interact with environmental risk factors and other genes implicated in AD and BIP.

Our results should be interpreted in light of the following limitations. We can neither exclude that some of the patients with AD have had BIP, nor that some of the patients with BIP will develop AD, which could have confounded our results. However; this could not explain the finding in the conjunctional FDR analyses of one locus implicated in AD and BIP with opposite directions of effect. Due to linkage disequilibrium among SNPs, our findings do not necessarily reflect causal variants, or that the same causal variants are involved in both traits. Although we found indications of modest polygenic overlap using conditional QQ-plots (**Figure 1**), we only detected two genetic loci jointly associated with both AD and BIP (**Figure 2**). However, the observed enrichment suggests that more shared SNPs will be identified when GWAS sample sizes increase (Schork et al., 2016). Further, we have only assessed the shared common genetic variants between AD and BIP. Other genetic variations, like rare structural variants, are also shown to increase the risk of AD and BIP (Lord et al., 2014; Cruceanu et al., 2017). Lastly, most participants in the data used in our study are of European ancestry. The generalizability of our findings to samples dominated by participants of other ancestries is unknown.

CONCLUSION

We find polygenic overlap between AD and BIP and identify novel loci associated with each trait and jointly with both traits, providing new insights into their genetic architecture. The genes *MARK2* and *VAC14* jointly implicated in AD and BIP are previously described to be involved in neuronal migration, tau phosphorylation, and endosomal homeostasis. Further experimental studies should examine if our findings translate to altered levels of transcription, pathophysiological processes, clinical features, and treatment response to lithium among patients with AD and BIP.

AUTHOR CONTRIBUTIONS

OD, PF, SH, and OA designed the protocol of the study. OA and OS obtained funding. AS, OS, OF, SH, and YW conducted the analyses. OS, AS, AW, OF, YW, SH, and OA interpreted the results. OD, OS, and AS drafted the manuscript. All authors contributed with the further writing of the manuscript and approved the final manuscript.

FUNDING

The study was supported by grants from by NIH (NS057198 and EB00790), the Research Council of Norway (229129, 213837, 223273, 251134, and 226971), the South-East Norway Regional Health Authority (2013–123, 2016–064), and KG Jebsen Foundation (SKGJ-2011–36).

ACKNOWLEDGMENTS

We would like to thank the International Genomics of Alzheimer's Project (IGAP) for providing summary results data for these analyses. The investigators within IGAP contributed to the design and implementation of IGAP and/or provided data but did not participate in analysis or writing of this report. IGAP was made possible by the generous participation of the control subjects, the patients, and their families. The i–Select chips was funded by the French National Foundation on Alzheimer's disease and related disorders. EADI was supported by the LABEX (laboratory of excellence program investment for the future) DISTALZ grant, Inserm, Institut Pasteur de Lille, Université de Lille 2 and the Lille University Hospital. GERAD was supported by the Medical Research Council (Grant n° 503480), Alzheimer's Research UK (Grant n° 503176), the Wellcome Trust (Grant n° 082604/2/07/Z) and German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) grant n° 01GI0102, 01GI0711, 01GI0420. CHARGE was partly supported by the NIH/NIA grant R01 AG033193 and the NIA AG081220 and AGES contract N01–AG–12100, the NHLBI grant R01 HL105756, the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA grants: U01 AG032984, U24 AG021886, U01 AG016976, and the Alzheimer's Association grant ADGC–10–196728. We would like to thank the Psychiatric Genetic Consortium 2 Bipolar Disorder Working Group (PGC2-BIP) for providing summary statistics for the study. The authors are listed in the **Supplementary Methods 1.5**. We would also like to thank the Human Brain Transcriptome atlas for giving permission to use their figures on spatio-temporal cerebral expression of the genes identified in our study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2019.00220/full#supplementary-material

REFERENCES

Alazami, A. M., Patel, N., Shamseldin, H. E., Anazi, S., Al-dosari, M. S., Alzahrani, F., et al. (2015). Accelerating novel candidate gene discovery in neurogenetic disorders via whole-exome sequencing of prescreened multiplex consanguineous families. *Cell Rep.* 10, 148–161. doi: 10.1016/j.celrep.2014.12.015

Alfonso, S. I., Callender, J. A., Hooli, B., Antal, C. E., Mullin, K., Sherman, M. A., et al. (2016). Gain-of-function mutations in protein kinase Ca (PKCa) may promote synaptic defects in alzheimer's disease. *Sci. Signal.* 9, 1–7. doi: 10.1126/scisignal.aaf6209

Altshuler, D. L., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., et al. (2010). A map of human genome variation from population-scale sequencing. *Nature* 467, 1061–1073. doi: 10.1038/nature09534

Alvarez-López, M. J., Molina-Martínez, P., Castro-Freire, M., Cosín-Tomás, M., Cristòfol, R., Párrizas, M., et al. (2014). Rcor2 underexpression in senescent mice: a target for inflammaging? *J. Neuroinflamm.* 11, 1–10. doi: 10.1186/1742-2094-11-126

Alzheimer, A. (1907). Uber eine eigenartige Erkrankung der Hirnrinde. *Allg. Zeitschrift Fur Psychiatr. Und Phychish-Gerichtliche Medizin* 64, 146–148.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* 4th ed. Washington, DC: American Psychiatric Association.

Andreassen, O. A., Desikan, R. S., Wang, Y., Thompson, W. K., Schork, A. J., Zuber, V., et al. (2015). Abundant genetic overlap between blood lipids and immune-mediated diseases indicates shared molecular genetic mechanisms. *PLoS One* 10:e0128048. doi: 10.1371/journal.pone.0123057

Andreassen, O. A., Djurovic, S., Thompson, W. K., Schork, A. J., Kendler, K. S., O'Donovan, M. C., et al. (2013a). Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. *Am. J. Hum. Genet.* 92, 197–209. doi: 10.1016/j.ajhg.2013.01.001

Andreassen, O. A., Thompson, W. K., Schork, A. J., Ripke, S., Mattingsdal, M., Kelsoe, J. R., et al. (2013b). Improved detection of common variants associated with schizophrenia and bipolar disorder using pleiotropy-informed conditional false discovery rate. *PLoS Genet.* 9:e1005544. doi: 10.1371/journal.pgen.1003455

Antonio, D., Concetta, C., Marco, C., and Alessandro, S. (2015). Enrichment pathway analysis. the inflammatory genetic background in bipolar disorder. . *J. Affect. Disord.* 179, 88–94. doi: 10.1016/j.jad.2015.03.032

Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., Duncan, L., et al. (2018). Analysis of shared heritability in common disorders of the brain. *Science* 360:6395. doi: 10.1126/science.aap8757

Baldessarini, R. J., Bolzani, L., Cruz, N., Jones, P. B., Lai, M., Lepri, B., et al. (2010). Onset-age of bipolar disorders at six international sites. *J. Affect. Disord.* 121, 143–146. doi: 10.1016/j.jad.2009.05.030

Balklava, Z., Niehage, C., Currinn, H., Mellor, L., Guscott, B., Poulin, G., et al. (2015). The amyloid precursor protein controls PIKfyve function. *PLoS One* 10:e0130485. doi: 10.1371/journal.pone.0130485

Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P.-R., et al. (2015). An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* 47, 1236–1241. doi: 10.1038/ng.3406

Cheng, C., Zandi, P., Stuart, E., Lin, C., Su, P., Alexander, G., et al. (2017). Association between lithium use and risk of alzheimer's disease. *J. Clin. Psychiatry* 78, e139–e145. doi: 10.4088/JCP.15m10304

Clark, E. A., Leach, K. L., Trojanowski, J. Q., and Lee, V. M. (1991). Characterization and differential distribution of the three major human protein kinase C isozymes (PKC alpha, PKC beta, and PKC gamma) of the central nervous system in normal and alzheimer's disease brains. *Lab. Invest.* 64, 35–44.

Colpo, G. D., Leffa, D. D., Köhler, C. A., Kapczinski, F., Quevedo, J., and Carvalho, A. F. (2015). Is bipolar disorder associated with accelerating aging? A meta-analysis of telomere length studies. *J. Affect. Disord.* 186, 241–248. doi: 10.1016/j.jad.2015.06.034

Corvin, A., Craddock, N., and Sullivan, P. F. (2010). Genome-wide association studies: a primer. *Psychol. Med.* 40, 1063–1077. doi: 10.1017/S0033291709991723

Cruceanu, C., Schmouth, J.-F., Torres-Platas, S. G., Lopez, J. P., Ambalavanan, A., Darcq, E., et al. (2017). Rare susceptibility variants for bipolar disorder suggest a role for G protein-coupled receptors. *Mol. Psychiatry* 23, 2050–2056. doi: 10.1038/mp.2017.223

Dannlowski, U., Kugel, H., Grotegerd, D., Redlich, R., Suchy, J., Opel, N., et al. (2015). NCAN cross-disorder risk variant is associated with limbic gray matter deficits in healthy subjects and major depression. *Neuropsychopharmacology* 40, 2510–2516. doi: 10.1038/npp.2015.86

Darrow, S. M., Verhoeven, J. E., Révész, D., Lindqvist, D., Penninx, B. W., Delucchi, K. L., et al. (2016). The association between psychiatric disorders and telomere length: a Meta-analysis involving 14,827 persons. *Psychosom. Med.* 78, 776–787. doi: 10.1097/PSY.0000000000000356.The

de Bakker, P. I. W., and Raychaudhuri, S. (2012). Interrogating the major histocompatibility complex with high-throughput genomics. *Hum. Mol. Genet.* 21, 29–36. doi: 10.1093/hmg/dds384

Delaneau, O., Marchini, J., and Zagury, J.-F. (2012). A linear complexity phasing method for thousands of genomes. *Nat. Methods* 9, 179–181. doi: 10.1038/nmeth.1785

Di Paolo, G., and De Camilli, P. (2006). Phosphoinositides in cell regulation and membrane dynamics. *Nature* 443, 651–657. doi: 10.1038/nature05185

Diazgranados, N., Kronstein, P., Khalife, S., Kammerer, W. A., Manji, H. K., and Zarate, C. A. (2010). A randomized add-on trial of an N-methyl-D-aspartateantagonist in treatment-resistant bipolar depression. *Arch. Gen. Psychiatry* 67, 793–802. doi: 10.1001/archgenpsychiatry.2010.90

Diniz, B. S., Teixeira, A. L., Cao, F., Gildengers, A., Soares, J. C., Butters, M. A., et al. (2017). History of bipolar disorder and the risk of dementia: a systematic review and meta-analysis. *Am. J. Geriatr. Psychiatry* 25, 357–362. doi: 10.1016/j.jagp.2016.11.014

Djurovic, S., Gustafsson, O., Mattingsdal, M., Athanasiu, L., Bjella, T., Tesli, M., et al. (2010). A genome-wide association study of bipolar disorder in Norwegian individuals, followed by replication in icelandic sample. *J. Affect. Disord.* 126, 312–316. doi: 10.1016/j.jad.2010.04.007

Dreses-Werringloer, U., Lambert, J., Vingtdeux, V., Zhao, H., Siebert, A., Jain, A., et al. (2008). A polymorphism in CALHM1 influences Ca2+ homeostasis, aβlevels, and alzheimer's disease risk. *Cell* 133, 1149–1161. doi: 10.1016/j.cell.2008.05.048

Einarsdottir, E., Peyrard-Janvid, M., Darki, F., Tuulari, J. J., Merisaari, H., Karlsson, L., et al. (2017). Identification of NCAN as a candidate gene for developmental dyslexia. *Sci. Rep.* 7, 1–11. doi: 10.1038/s41598-017-10175-10177

Escott-Price, V., Sims, R., Bannister, C., Harold, D., Vronskaya, M., Majounie, E., et al. (2015). Common polygenic variation enhances risk prediction for alzheimer's disease. *Brain* 138, 3673–3684. doi: 10.1093/brain/awv268

Forlenza, O., Aprahamian, I., Radanovic, M., Talib, L., Camargo, M., Stella, F., et al. (2016). Cognitive impairment in late-life bipolar disorder is not associated with Alzheimer's disease pathological signature in the cerebrospinal fluid. *Bipolar Disord.* 18, 63–70. doi: 10.1111/bdi.12360

Frei, O., Holland, D., Smeland, O. B., Shadrin, A. A., Chieh, C. F., Maeland, S., et al. (2018). Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *bioRxiv* [Preprint]. doi: 10.1101/240275

Freland, L., and Beaulieu, J.-M. (2012). Inhibition of GSK3 by lithium, from single molecules to signaling networks. *Front. Mol. Neurosci.* 5:14. doi: 10.3389/fnmol.2012.00014

Gatz, M., Reynolds, C. A., Fratiglioni, L., Johansson, B., Mortimer, J. A., Berg, S., et al. (2006). Role of genes and environments for explaining alzheimer disease. *Arch. Gen. Psychiatry* 63, 168–174. doi: 10.1001/archpsyc.63.2.168

Gerhard, T., Devanand, D. P., Huang, C., Crystal, S., and Olfson, M. (2015). Lithium treatment and risk for dementia in adults with bipolar disorder: population-based cohort study. *Br. J. Psychiatry* 207, 46–51. doi: 10.1192/bjp.bp.114.154047

Godefroy, O., Bakchine, S., Verny, M., Delabrousse-Mayoux, J. P., and Roussel, M. (2016). Characteristics of alzheimer's disease patients with severe executive disorders. *J. Alzheimer's Dis.* 51, 815–825. doi: 10.3233/JAD-150971

Gold, C. A., and Budson, A. E. (2008). Memory loss in Alzheimer's disease: implications for development of therapeutics. *Expert Rev. Neurother.* 8, 1879–1891. doi: 10.1586/14737175.8.12.1879

Goldstein, B. I., Kemp, D. E., Soczynska, J. K., and McIntyre, R. S. (2009). Inflammation and the phenomenology, pathophysiology, comorbidity, and treatment of bipolar disorder: a systematic review of the literature. *J. Clin. Psychiatry* 70, 1078–1090. doi: 10.4088/JCP.08r04505

Grande, I., Berk, M., Birmaher, B., and Vieta, E. (2016). Bipolar disorder. *Lancet* 387, 1561–1572. doi: 10.1016/S0140-6736(15)00241-X

Gu, G. J., Wu, D., Lund, H., Sunnemark, D., Kvist, A. J., Milner, R., et al. (2013). Elevated MARK2-dependent phosphorylation of tau in alzheimer's disease. *J. Alzheimer's Dis.* 33, 699–713. doi: 10.3233/JAD-2012-121357

Gu, P., Deng, W., Lei, M., and Chang, S. (2013). Single strand DNA binding proteins 1 and 2 protect newly replicated telomeres. *Cell Res.* 23, 705–719. doi: 10.1038/cr.2013.31

Gueneau, L., Fish, R. J., Shamseldin, H. E., Voisin, N., Mau-Them, F. T., Preiksaitiene, E., et al. (2018). KIAA1109 variants are associated with a severe disorder of brain development and arthrogryposis. *Am. J. Hum. Genet.* 102, 116–132. doi: 10.1016/j.ajhg.2017.12.002

Gulaj, E., Pawlak, K., Bien, B., and Pawlak, D. (2010). Kynurenine and its metabolites in alzheimer's disease patients. *Adv. Med. Sci.* 55, 204–211. doi: 10.2478/v10039-010-0023-6

Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., Hamshere, M. L., et al. (2009). Genome-wide association study identifies variants at CLU and PICALM associated with alzheimer's disease. *Nat. Genet.* 41, 1088–1093. doi: 10.1038/ng.440

Heath, S. C., Gut, I. G., Brennan, P., McKay, J. D., Bencko, V., Fabianova, E., et al. (2008). Investigation of the fine structure of european populations with applications to disease association studies. *Eur. J. Hum. Genet.* 16, 1413–1429. doi: 10.1038/ejhg.2008.210

Heneka, M. T., Carson, M. J., Khoury, J., El Landreth, G. E., Brosseron, F., Feinstein, D. L., et al. (2015). Neuroinflammation in alzheimer's disease. *Lancet Neurol.* 14, 388–405. doi: 10.1016/S1474-4422(15)70016-70015

Howard, R., McShane, R., Lindesay, J., Ritchie, C., Baldwin, A., Barber, R., et al. (2012). Donepezil and memantine for moderate-to-severe alzheimer's disease. *N. Engl. J. Med.* 366, 893–903. doi: 10.1056/NEJMoa1106668

Howie, B. N., Donnelly, P., and Marchini, J. (2009). A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 5:e1000529. doi: 10.1371/journal.pgen.1000529

Itokawa, M., Yamada, K., Iwayama-Shigeno, Y., Ishitsuka, Y., Detera-Wadleigh, S., and Yoshikawa, T. (2003). Genetic analysis of a functional GRIN2A promoter (GT)n repeat in bipolar disorder pedigrees in humans. *Neurosci. Lett.* 345, 53–56. doi: 10.1016/S0304-3940(03)00501-509

Jack, C. R., Knopman, D. S., Jagust, W. J., Petersen, R. C., Weiner, M. W., Aisen, P. S., et al. (2013). Tracking pathophysiological processes in alzheimer's disease: An updated hypothetical model of dynamic biomarkers. *Lancet Neurol.* 12, 207–216. doi: 10.1016/S1474-4422(12)70291-70290

Jakobsson, J., Palsson, E., Sellgren, C., Rydberg, F., Ekman, A., Zetterberg, H., et al. (2016). CACNA1C polymorphism and altered phosphorylation of tau in bipolar disorder. *Br. J. Psychiatry* 208, 195–196. doi: 10.1192/bjp.bp.114.159806

Jakobsson, J., Zetterberg, H., Blennow, K., Johan Ekman, C., Johansson, A. G. M., and Landén, M. (2013). Altered concentrations of amyloid precursor protein metabolites in the cerebrospinal fluid of patients with bipolar disorder. *Neuropsychopharmacology* 38, 664–672. doi: 10.1038/npp.2012.231

Jiang, X., Detera-Wadleigh, S. D., Akula, N., Mallon, B. S., Hou, L., Xiao, T., et al. (2018). Sodium valproate rescues expression of TRANK1 in iPSC-derived neural cells that carry a genetic variant associated with serious mental illness. *Mol. Psychiatry* doi: 10.1038/s41380-018-0207-201 [Epub ahead of print].

Jun, G., Naj, A. C., Beecham, G. W., Wang, L. S., Buros, J., Gallins, P. J., et al. (2010). Meta-analysis confirms CR1, CLU, and PICALM as Alzheimer disease risk loci and reveals interactions with APOE genotypes. *Arch. Neurol.* 67, 1473–1484. doi: 10.1001/archneurol.2010.201

Kang, H. J., Kawasawa, Y. I., Cheng, F., Zhu, Y., Xu, X., Li, M., et al. (2011). Spatio-temporal transcriptome of the human brain. *Nature* 478, 483–489. doi: 10.1038/nature10523

Kessing, L. V., Forman, J. L., and Andersen, P. K. (2010). Does lithium protect against dementia? *Bipolar Disord.* 12, 87–94. doi: 10.1111/j.1399-5618.2009.00788.x

Kessing, L. V., and Nilsson, F. M. (2003). Increased risk of developing dementia in patients with major affective disorders compared to patients with other medical illnesses. *J. Affect. Disord.* 73, 261–269. doi: 10.1016/S0165-0327(02)00004-6

Kessing, L. V., Olsen, E. W., Mortensen, P. B., and Andersen, P. K. (1999). Dementia in affective disorder: a case-register study. *Acta Psychiatr. Scand.* 100, 176–185. doi: 10.1111/j.1600-0447.1999.tb10843.x

Kieseppä, T., and Partonen, T. (2004). High concordance of bipolar I disorder in a nationwide sample of twins. *Am. J. Psychiatry* 161, 1814–1821. doi: 10.1176/appi.ajp.161.10.1814

Kittel-Schneider, S., Lorenz, C., Auer, J., Weiß, L., and Reif, A. (2016). DGKH genetic risk variant influences gene expression in bipolar affective disorder. *J. Affect. Disord.* 198, 148–157. doi: 10.1016/j.jad.2016.03.041

Knowles, E. E. M., Meikle, P. J., Huynh, K., Göring, H. H. H., Olvera, R. L., Mathias, S. R., et al. (2017). Serum phosphatidylinositol as a biomarker for bipolar disorder liability. *Bipolar Disord.* 19, 107–115. doi: 10.1111/bdi.12468

Koedam, E. L. G. E., Lauffer, V., Van Der Vlies, A. E., Van Der Flier, W. M., Scheltens, P., and Pijnenburg, Y. A. L. (2010). Early-versus late-onset alzheimer's disease: more than age alone. *J. Alzheimer's Dis.* 19, 1401–1408. doi: 10.3233/JAD-2010-1337

Kon, E., Cossard, A., and Jossin, Y. (2017). Neuronal polarity in the embryonic mammalian cerebral cortex. *Front. Cell. Neurosci.* 11:163. doi: 10.3389/fncel.2017.00163

Kosuga, S., Tashiro, E., Kajioka, T., Ueki, M., Shimizu, Y., and Imoto, M. (2005). GSK-3β directly phosphorylates and activates MARK2/PAR-1. *J. Biol. Chem.* 280, 42715–42722. doi: 10.1074/jbc.M507941200

Kraepelin, E. (1921). *Manic-Depressive Insanity and Paranoia*. Bristol: Thoemmes Press.

Lambert, J.-C., Grenier-Boley, B., Chouraki, V., Heath, S., Zelenika, D., Fievet, N., et al. (2010). Implication of the immune system in alzheimer's disease: evidence from genome-wide pathway analysis. *J. Alzheimer's Dis.* 20, 1107–1118. doi: 10.3233/JAD-2010-100018

Lambert, J.-C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., et al. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for alzheimer's disease. *Nat. Genet.* 45, 1452–1458. doi: 10.1038/ng.2802

Lee, S. H., Harold, D., Nyholt, D. R., Goddard, M. E., Zondervan, K. T., Williams, J., et al. (2013). Estimation and partitioning of polygenic variation captured by common SNPs for alzheimer's disease, multiple sclerosis and endometriosis. *Hum. Mol. Genet.* 22, 832–841. doi: 10.1093/hmg/dds491

Lee, S. H., Wray, N. R., Goddard, M. E., and Visscher, P. M. (2011). Estimating missing heritability for disease from genome-wide association studies. *Am. J. Hum. Genet.* 88, 294–305. doi: 10.1016/j.ajhg.2011.02.002

Li, F., and Tsien, J. (2009). Clinical implications of basic research: memory and the NMDA receptors. *N. Engl. J. Med.* 361, 302–303. doi: 10.1056/NEJMcibr0902052

Li, Y., Willer, C. J., Ding, J., Scheet, P., and Abecasis, G. R. (2010). MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.* 34, 816–834. doi: 10.1002/gepi.20533

Lichtenstein, P., Yip, B. H., Björk, C., Pawitan, Y., Cannon, T. D., Sullivan, P. F., et al. (2009). Common genetic determinants of schizophrenia and bipolar disorder in swedish families: a population-based study. *Lancet* 373, 234–239. doi: 10.1016/S0140-6736(09)60072-60076

Lord, J., Lu, A. J., and Cruchaga, C. (2014). Identification of rare variants in alzheimer's disease. *Front. Genet.* 5:369. doi: 10.3389/fgene.2014.00369

Machiela, M. J., and Chanock, S. J. (2015). Genetics and population analysis LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* 31, 3555–3557. doi: 10.1093/bioinformatics/btv402

MacLeod, C. A., and Donaldson, D. I. (2014). PRKCA polymorphism changes the neural basis of episodic remembering in healthy individuals. *PLoS One* 9:e98018. doi: 10.1371/journal.pone.0098018

Macritchie, K. A., Geddes, J. R., Scott, J., Haslam, D. R., and Goodwin, G. M. (2001). Valproic acid, valproate and divalproex in the maintenance treatment of bipolar disorder. *Cochrane Database Syst. Rev.* 3:CD003196. doi: 10.1002/14651858.CD003196

Maddison, D. C., and Giorgini, F. (2015). The kynurenine pathway and neurodegenerative disease. *Semin. Cell Dev. Biol.* 40, 134–141. doi: 10.1016/j.semcdb.2015.03.002

Maidment, I. D., Fox, C. G., Boustani, M., Rodriguez, J., Brown, R. C., and Katona, C. L. (2008). Efficacy of memantine on behavioral and psychological symptoms related to dementia: a systematic meta-analysis. *Ann. Pharmacother.* 42, 32–38. doi: 10.1345/aph.1K372

Martino, D. J., Samamé, C., Ibañez, A., and Strejilevich, S. A. (2015). Neurocognitive functioning in the premorbid stage and in the first episode of bipolar disorder: a systematic review. *Psychiatry Res.* 226, 23–30. doi: 10.1016/j.psychres.2014.12.044

Matenia, D., and Mandelkow, E. M. (2009). The tau of MARK: a polarized view of the cytoskeleton. *Trends Biochem. Sci.* 34, 332–342. doi: 10.1016/j.tibs.2009.03.008

Matsunaga, S., Kishi, T., Annas, P., Basun, H., Hampel, H., and Iwata, N. (2015). Lithium as a treatment for alzheimer's disease: a systematic review and meta-analysis. *J. Alzheimer's Dis.* 48, 403–410. doi: 10.3233/JAD-150437

McCartney, A. J., Zhang, Y., and Weisman, L. S. (2014). Phosphatidylinositol 3,5-bisphosphate: low abundance, high significance. *Bioessays* 36, 52–64. doi: 10.1002/bies.201300012

McGuffin, P., Rijsdijk, F., Andrew, M., Sham, P., Katz, R., and Cardno, A. (2003). The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Arch. Gen. Psychiatry* 60, 497–502. doi: 10.1001/archpsyc.60.5.497

McKhann, G., Drachman, D., and Folstein, M. (1984). Clinical diagnosis of alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of department of health and human services task force on alzheimer's disease. *Neurology* 34, 939–944. doi: 10.1212/WNL.34.7.939

Miller, C. L., Llenos, I. C., Dulay, J. R., and Weis, S. (2006). Upregulation of the initiating step of the kynurenine pathway in postmortem anterior cingulate cortex from individuals with schizophrenia and bipolar disorder. *Brain Res.* 1073–1074, 25–37. doi: 10.1016/j.brainres.2005.12.056

Miranda, A. M., Lasiecka, Z. M., Xu, Y., Neufeld, J., Shahriar, S., Simoes, S., et al. (2018). Neuronal lysosomal dysfunction releases exosomes harboring APP C-terminal fragments and unique lipid signatures. *Nat. Commun.* 9:291. doi: 10.1038/s41467-017-02533-w

Myint, A. M., Kim, Y. K., Verkerk, R., Park, S. H., Scharpé, S., Steinbusch, H. W. M., et al. (2007). Tryptophan breakdown pathway in bipolar mania. *J. Affect. Disord.* 102, 65–72. doi: 10.1016/j.jad.2006.12.008

Newell, K. L., Hyman, B. T., Growdon, J. H., and Hedley-Whyte, E. T. (1999). Application of the national institute on aging (NIA)-reagan institute criteria for the neuropathological diagnosis of alzheimer disease. *J. Neuropathol. Exp. Neurol.* 58, 1147–1155. doi: 10.1097/00005072-199911000-199911004

Nunes, P. V., Forlenxa, O. V., and Gattaz, W. F. (2007). Lithium and risk for alzheimer's disease in elderly patients with bipolar disorder. *Br. J. Psychiatry* 190, 359–361. doi: 10.1192/bjp.bp.106.029868

O'Brien, R. J., and Wong, P. C. (2011). Amyloid precursor protein processing and alzheimer's disease. *Annu. Rev. Neurosci.* 34, 185–204. doi: 10.1146/annurev-neuro-061010-113613

Okamoto, N., Tsuchiya, Y., Miya, F., Tsunoda, T., Yamashita, K., and Boroevich, K. A. (2017). A novel genetic syndrome with STARD9 mutation and abnormal spindle morphology. *Am. J. Med. Genet.* 173, 2690–2696. doi: 10.1002/ajmg.a.38391

O'Shea, K. S., and McInnis, M. G. (2016). Neurodevelopmental origins of bipolar disorder: iPSC models. *Mol. Cell. Neurosci.* 73, 63–83. doi: 10.1016/j.mcn.2015.11.006

Pagnamenta, A. T., Bacchelli, E., De Jonge, M. V., Mirza, G., Scerri, T. S., Minopoli, F., et al. (2010). Characterization of a family with rare deletions in CNTNAP5 and DOCK4 suggests novel risk loci for autism and dyslexia. *Biol. Psychiatry* 68, 320–328. doi: 10.1016/j.biopsych.2010.02.002

Powell, T. R., Dima, D., Frangou, S., and Breen, G. (2017). Telomere length and bipolar disorder. *Neuropsychopharmacology* 43, 445–453. doi: 10.1038/npp.2017.125

Psaty, B. M., O'Donnell, C. J., Gudnason, V., Lunetta, K. L., Folsom, A. R., Rotter, J. I., et al. (2009). Cohorts for heart and aging research in genomic epidemiology (CHARGE) consortium design of prospective meta-analyses of genome-wide

association studies from 5 cohorts. *Circ. Cardiovasc. Genet.* 2, 73–80. doi: 10.1161/CIRCGENETICS.108.829747

Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., et al. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460, 748–752. doi: 10.1038/nature08185

Rahman, A., Ting, K., Cullen, K. M., Braidy, N., Brew, B. J., and Guillemin, G. J. (2009). The excitotoxin quinolinic acid induces tau phosphorylation in human neurons. *PLoS One* 4:e6344. doi: 10.1371/journal.pone.0006344

Raum, H., Dietsche, B., Nagels, A., Witt, S. H., Rietschel, M., Kircher, T., et al. (2015). A genome-wide supported psychiatric risk variant in NCAN influences brain function and cognitive performance in healthy subjects. *Hum. Brain Mapp.* 36, 378–390. doi: 10.1002/hbm.22635

Reiner, O., Shmueli, A., and Sapir, T. (2009). Neuronal migration and neurodegeneration: 2 Sides of the same coin. *Cereb. Cortex* 19, i42–i48. doi: 10.1093/cercor/bhp039

Reisberg, B., Doody, R., Stöffner, A., Schmitt, F., Ferris, S., Möbius, H. J., et al. (2003). Memantine in moderate-to-severe alzheimer's disease. *N. Engl. J. Med.* 348, 1333–1341. doi: 10.1056/nejm200308073490616

Ridge, P. G., Hoyt, K. B., Boehme, K., Mukherjee, S., Crane, P. K., Haines, J. L., et al. (2016). Assessment of the genetic variance of late-onset alzheimer's disease. *Neurobiol. Aging* 41, 200.e13–200.e20. doi: 10.1016/j.neurobiolaging.2016.02.024

Ridge, P. G., Mukherjee, S., Crane, P. K., and Kauwe, J. S. K. (2013). Alzheimer's disease: analyzing the missing heritability. *PLoS One* 8:e79771. doi: 10.1371/journal.pone.0079771

Sanches, M., Keshavan, M. S., Brambilla, P., and Soares, J. C. (2008). Neurodevelopmental basis of bipolar disorder: A critical appraisal. *Prog. Neuro Psychopharmacol. Biol. Psychiatry* 32, 1617–1627. doi: 10.1016/j.pnpbp.2008.04.017

Savitz, J., Dantzer, R., Wurfel, B. E., Victor, T. A., Ford, B. N., Bodurka, J., et al. (2015). Neuroprotective kynurenine metabolite indices are abnormally reduced and positively associated with hippocampal and amygdalar volume in bipolar disorder. *Psychoneuroendocrinology* 52, 200–211. doi: 10.1016/j.psyneuen.2014.11.015

Scheltens, P., Blennow, K., Breteler, M. M. B., de Strooper, B., Frisoni, G. B., Salloway, S., et al. (2016). Alzheimer's disease. *Lancet* 388, 505–517. doi: 10.1016/S0140-6736(15)01124-1121

Schork, A. J., Wang, Y., Thompson, W. K., Dale, A. M., and Andreassen, O. A. (2016). New statistical approaches exploit the polygenic architecture of schizophrenia - implications for the underlying neurobiology. *Curr. Opin. Neurobiol.* 36, 89–98. doi: 10.1016/j.conb.2015.10.008

Schott, J. M., Crutch, S. J., Carrasquillo, M. M., Uphill, J., Shakespeare, T. J., Ryan, N. S., et al. (2016). Genetic risk factors for the posterior cortical atrophy variant of alzheimer's disease. *Alzheimer's Dement.* 12, 862–871. doi: 10.1016/j.jalz.2016.01.010

Smeland, O. B., Bahrami, S., Frei, O., Shadrin, A., O'Connell, K., Savage, J., et al. (2019). Genome-wide analysis reveals extensive genetic overlap between schizophrenia, bipolar disorder, and intelligence. *Mol. Psychiatry* doi: 10.1038/s41380-018-0332-x [Epub ahead of print].

Smeland, O. B., Frei, O., Kauppi, K., Hill, W. D., Li, W., Wang, Y., et al. (2017). Identification of genetic loci jointly influencing schizophrenia risk and the cognitive traits of verbal-numerical reasoning, reaction time, and general cognitive function. *JAMA Psychiatry* 74, 1065–1075. doi: 10.1001/jamapsychiatry.2017.1986

Stahl, E., Breen, G., Forstner, A., McQuillin, A., Ripke, S., Bipolar Disorder Working Group of the Psychiatric Genomics Consortium, et al. (2019). Genomewide association study identifies 30 loci associated with bipolar disorder. *bioRxiv* [Preprint]. doi: 10.1101/173062

Timm, T., Balusamy, K., Li, X., Biernat, J., Mandelkow, E., and Mandelkow, E. M. (2008). Glycogen Synthase Kinase (GSK) 3β directly phosphorylates serine 212 in the regulatory loop and inhibits microtubule affinity-regulating kinase (MARK) 2. *J. Biol. Chem.* 283, 18873–18882. doi: 10.1074/jbc.M706596200

Torres, I. J., Boudreau, V. G., and Yatham, L. N. (2007). Neuropsychological functioning in euthymic bipolar disorder: a meta-analysis. *Acta Psychiatr. Scand.* 116, 17–26. doi: 10.1111/j.1600-0447.2007.01055.x

Traut, W., Weichenhan, D., Himmelbauer, H., and Winking, H. (2006). New members of the neurexin superfamily: Multiple rodent homologues of the human CASPR5 gene. *Mamm. Genome* 17, 723–731. doi: 10.1007/s00335-005-0157-151

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., Mccarthy, M. I., Brown, M. A., et al. (2017). 10 years of GWAS discovery: biology, function, and translation. *Am. J. Hum. Genet.* 101, 5–22. doi: 10.1016/j.ajhg.2017.06.005

Vos, T., Allen, C., Arora, M., Barber, R. M., Brown, A., Carter, A., et al. (2016). Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the global burden of disease study 2015. *Lancet* 388, 1545–1602. doi: 10.1016/S0140-6736(16)31678-31676

Wang, H.-Y., Markowitz, P., Levinson, D., Undie, A. S., and Friedman, E. (1999). Increased membrane-associated protein kinase C activity and translocation in blood platelets from bipolar affective disorder patients. *J. Psychiatr. Res.* 33, 171–179. doi: 10.1016/S0022-3956(98)90057-90057

Wang, Y., Thompson, W. K., Schork, A. J., Holland, D., Chen, C.-H., Bettella, F., et al. (2016a). Leveraging genomic annotations and pleiotropic enrichment for improved replication rates in schizophrenia GWAS. *PLoS Genet.* 12:e1005803. doi: 10.1371/journal.pgen.1005803

Wang, Y., Wu, Q., Yang, P., Wang, C., Liu, J., Ding, W., et al. (2016b). LSD1 co-repressor Rcor2 orchestrates neurogenesis in the developing mouse brain. *Nat. Commun.* 7, 1–14. doi: 10.1038/ncomms10481

World Health Organization. (1977). *International Classification of Diseases, Ninth Revision*. Geneva: World Health Organization.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.

Yu, H., Yan, H., Wang, L., Li, J., Tan, L., Deng, W., et al. (2018). Five novel loci associated with antipsychotic treatment response in patients with schizophrenia: a genome-wide association study. *Lancet Psychiatry* 5, 327–338. doi: 10.1016/S2215-0366(18)30049-X

Yuan, H., Low, C.-M., Moody, O. A., Jenkins, A., and Traynelis, S. F. (2015). Ionotropic GABA and glutamate receptor mutations and human neurologic diseases. *Mol. Pharmacol.* 88, 203–217. doi: 10.1124/mol.115.097998

Zarate, C. Jr., Brutsche, N., Ibrahim, L., Franco-Chaves, J., Diazgranados, N., Cravchik, A., et al. (2012). Replication of ketamine's antidepressant efficacy in bipolar depression: a randomized controlled add-on trial. *Biol. Psychiatry* 71, 939–946. doi: 10.1016/j.biopsych.2011.12.010

Zhan, Y., Song, C., Karlsson, R., Tillander, A., Reynolds, C. A., Pedersen, N. L., et al. (2015). Telomere length shortening and alzheimer disease - a mendelian randomization study. *JAMA Neurol.* 72, 1202–1203. doi: 10.1001/jamaneurol.2015.1513

Zhang, Y., Zolov, S. N., Chow, C. Y., Slutsky, S. G., Richardson, S. C., Piper, R. C., et al. (2007). Loss of Vac14, a regulator of the signaling lipid phosphatidylinositol 3,5-bisphosphate, results in neurodegeneration in mice. *Proc. Natl. Acad. Sci.* 104, 17518–17523. doi: 10.1073/pnas.0702275104

Conflict of Interest Statement: OA has received a speaker's honorarium from Lundbeck and has a patent application (US 20150356243) pending. AD also applied for this patent application and assigned it to UC San Diego. AD has additional disclosures outside the present work: founder, equity holder, and advisory board member for CorTechs Labs, advisory board member of Human Longevity, recipient of non-financial research support from General Electric Healthcare.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership