Neuroscience Institute Publications   Western Institute for Neuroscience

8-1-2022

Targeted copy number variant identification across the neurodegenerative disease spectrum

Allison A. Dilliott
Institut-Hôpital Neurologique de Montréal

Kristina K. Zhang
Schulich School of Medicine & Dentistry

Jian Wang
Robarts Research Institute

Agessandro Abrahao
University of Toronto

Malcolm A. Binns
Rotman Research Institute

See next page for additional authors

Follow this and additional works at: https://ir.lib.uwo.ca/neurosci_inst_pubs

Authors

Allison A. Dilliott, Kristina K. Zhang, Jian Wang, Agessandro Abrahao, Malcolm A. Binns, Sandra E. Black, Michael Borrie, Dar Dowlatshahi, Elizabeth Finger, Corinne E. Fischer, Andrew Frank, Morris Freedman, David Grimes, Ayman Hassan, Mandar Jog, Sanjeev Kumar, Anthony E. Lang, Jennifer Mandzia, Mario Masellis, Stephen H. Pasternak, Bruce G. Pollock, Tarek K. Rajji, Ekaterina Rogaeva, Demetrios J. Sahlas, Gustavo Saposnik, Christine Sato, Dallas Seitz, Christen Shoesmith, Thomas D.L. Steeves, and Richard H. Swartz

Molecular Genetics & Genomic Medicine
Open Access **WILEY**

ORIGINAL ARTICLE

Targeted copy number variant identification across the neurodegenerative disease spectrum

Allison A. Dilliott[1] | **Kristina K. Zhang**[2] | **Jian Wang**[3] | **Agessandro Abrahao**[4] | **Malcolm A. Binns**[5,6] | **Sandra E. Black**[4,7] | **Michael Borrie**[8,9] | **Dar Dowlatshahi**[10,11] | **Elizabeth Finger**[12,13] | **Corinne E. Fischer**[14] | **Andrew Frank**[10,15] | **Morris Freedman**[5,16] | **David Grimes**[10,11] | **Ayman Hassan**[17] | **Mandar Jog**[12,18] | **Sanjeev Kumar**[19,20] | **Anthony E. Lang**[21,22] | **Jennifer Mandzia**[12] | **Mario Masellis**[22,23] | **Stephen H. Pasternak**[3,12,24] | **Bruce G. Pollock**[19,20] | **Tarek K. Rajji**[19,25] | **Ekaterina Rogaeva**[26] | **Demetrios J. Sahlas**[27] | **Gustavo Saposnik**[28,29] | **Christine Sato**[26] | **Dallas Seitz**[30] | **Christen Shoesmith**[18] | **Thomas D. L. Steeves**[12,31] | **Richard H. Swartz**[4,22,7] | **Brian Tan**[5] | **David F. Tang-Wai**[22,32,33] | **Maria C. Tartaglia**[26] | **John Turnbull**[27] | **Lorne Zinman**[4] | **ONDRI Investigators** | **Robert A. Hegele**[3,34]

[1]Department of Neurology and Neurosurgery, Montreal Neurological Institute and Hospital, McGill University, Montréal, Quebec, Canada

[2]Department of Microbiology & Immunology, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada

[3]Robarts Research Institute, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada

[4]Division of Neurology, Department of Medicine, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Ontario, Canada

[5]Rotman Research Institute, Baycrest Health Sciences, Toronto, Ontario, Canada

[6]Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada

[7]LCCampbell Cognitive Neurology Research Unit, Hurvitz Brain Sciences Research Program Sunnybrook Health Sciences Research Program, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

[8]St. Joseph's Health Care Centre, London, Ontario, Canada

[9]Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada

[10]Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada

[11]Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

[12]Department of Clinical Neurological Sciences, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada

[13]Lawson Health Research Institute, London, Ontario, Canada

[14]Keenan Research Centre for Biomedical Science, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Ontario, Canada

[15]Bruyère Research Institute, Ottawa, Ontario, Canada

[16]Division of Neurology, Department of Medicine, Baycrest Health Sciences, Mt. Sinai Hospital, University of Toronto, Toronto, Ontario, Canada

[17]Thunder Bay Regional Research Institute, Northern Ontario School of Medicine, Thunder Bay, Ontario, Canada

[18]London Health Sciences Centre, London, Ontario, Canada

[19]Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, Ontario, Canada

[20]Department of Psychiatry, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada

[21]Edmond J. Safra Program in Parkinson's Disease and the Morton and Gloria Shulman Movement Disorders Clinic, Toronto Western Hospital, Toronto, Ontario, Canada

[22]Department of Medicine, Division of Neurology, University of Toronto, Toronto, Ontario, Canada

[23]Cognitive & Movement Disorders Clinic, L.C. Campbell Cognitive Neurology Research Unit, Hurvitz Brain ScienceProgram, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada

[24]Cognitive Neurology and Alzheimer's Disease Research Centre, Parkwood Institute, St. Joseph's Health Care, London, Ontario, Canada

[25]Department of Psychiatry and Toronto Dementia Research Alliance, University of Toronto, Toronto, Ontario, Canada

[26]Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, Ontario, Canada

[27]Department of Medicine, McMaster University, Hamilton, Ontario, Canada

[28]Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, Ontario, Canada

[29]Clinical Outcomes and Decision Neuroscience Unit, St. Michael's Hospital, University of Toronto, Toronto, Ontario, Canada

[30]Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

[31]Division of Neurology, St. Michael's Hospital, Toronto, Ontario, Canada

[32]Krembil Research Institute, Toronto Western Hospital, Toronto, Ontario, Canada

[33]University Health Network Memory Clinic, Toronto Western Hospital, Toronto, Ontario, Canada

[34]Department of Biochemistry, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada

Correspondence
Robert A. Hegele, Robarts Research Institute, Western University, 4288A-1151 Richmond Street North, London, ON N6A 5C1, Canada.
Email: hegele@robarts.ca

Allison A. Dilliott, Montreal Neurological Institute and Hospital, McGill University, 3801 Rue University, Montréal, QC H3A 2B4, Canada.
Email: allison.dilliott@mcgill.ca

Abstract

Background: Although genetic factors are known to contribute to neurodegenerative disease susceptibility, there remains a large amount of heritability unaccounted for across the diagnoses. Copy number variants (CNVs) contribute to these phenotypes, but their presence and influence on disease state remains relatively understudied.

Methods: Here, we applied a depth of coverage approach to detect CNVs in 80 genes previously associated with neurodegenerative disease within participants of the Ontario Neurodegenerative Disease Research Initiative ($n = 519$).

Results: In total, we identified and validated four CNVs in the cohort, including: (1) a heterozygous deletion of exon 5 in *OPTN* in an Alzheimer's disease participant; (2) a duplication of exons 1–5 in *PARK7* in an amyotrophic lateral sclerosis participant; (3) a duplication of >3 Mb, which encompassed *ABCC6*, in a cerebrovascular disease (CVD) participant; and (4) a duplication of exons 7–11 in *SAMHD1* in a mild cognitive impairment participant. We also identified 43 additional CNVs that may be candidates for future replication studies.

Conclusion: The identification of the CNVs suggests a portion of the apparent missing heritability of the phenotypes may be due to these structural variants, and their assessment is imperative for a thorough understanding of the genetic spectrum of neurodegeneration.

KEYWORDS

cerebrovascular disease, copy number variants, neurodegenerative disease, next-generation sequencing

1 | INTRODUCTION

Neurodegenerative diseases are a collection of progressive conditions characterized by neuronal degeneration and protein aggregation within the brain. Although typically defining behavioral and/or cognitive phenotypes, such as Alzheimer's disease (AD), frontotemporal dementia (FTD), and mild cognitive impairment (MCI) or motor phenotypes, such as amyotrophic lateral sclerosis (ALS) and Parkinson's disease (PD), the term may also encompass neurodegenerative phenotypes that result from, or present alongside, cerebrovascular disease (CVD).

Genetic factors can increase one's risk of developing neurodegenerative disease considerably, with relatively high heritability estimates across the various diagnoses (Cacace et al., 2016; Greaves & Rohrer, 2019; Mejzini et al., 2019). However, our existing understanding of the genetic contributors to neurodegenerative disease fail to

reach these estimates, leaving a large amount of missing heritability (Cacace et al., 2016; Hagenaars et al., 2018; Keller et al., 2012; Mejzini et al., 2019). For example, while AD displays heritability estimates of ~70%, depending on age of disease onset, up to 90% of cases remain genetically unexplained (Cacace et al., 2016; Karlsson et al., 2022).

Copy number variants (CNVs) are large-scale deletions or duplications of DNA of at least 50 base pairs (bp) in length (Feuk et al., 2006). While CNVs are generally common across the genome (Redon et al., 2006), when they involve genes, these structural variants can have large phenotypic impacts, including affecting gene expression, organization, and dosage (Stranger et al., 2007). CNVs have been shown to contribute to neurodegenerative disease presentation, including in individuals with AD (Cuccaro et al., 2017; Ghani et al., 2012; Hooli et al., 2014); ALS (Morello et al., 2018); and PD (Nuytemans et al., 2010). Yet studies identifying these variants in neurodegenerative disease cohorts are relatively sparse, potentially due to the previous intricacies of accurately detecting CNVs. It is therefore hypothesized that some of the missing heritability of neurodegeneration may be accounted for by these large-scale variants.

Bioinformatics algorithms have recently made the identification of CNVs more accessible, with the ability to detect variants using next-generation sequencing (NGS) and a depth of coverage (DOC) assessment (Iacocca et al., 2017). Here, we leveraged this approach to identify CNVs across the participants of the Ontario Neurodegenerative Disease Research Initiative (ONDRI), a multi-cohort study aiming to characterize a selection of neurodegenerative diseases, including AD, ALS, FTD, MCI, and PD, as well as CVD and its potential influence on neurodegeneration. Previously, the ONDRI cohort was genetically sequenced using the ONDRISeq NGS targeted panel, which covers 80 genes known to contribute to neurodegenerative diseases (Dilliott et al., 2018; Farhan et al., 2016). Following the identification of CNVs using a DOC approach, we also aimed to validate a subset of CNVs using either breakpoint analysis or exome sequencing.

2 | MATERIALS AND METHODS

2.1 | Study participants and ethical compliance

ONDRI enrolled 520 individuals from clinical sites across Ontario who passed preliminary screening and were each clinically diagnosed with one of the following conditions: (1) AD; (2) ALS; (3) CVD; (4) FTD; (5) MCI; or (6) PD. Research ethics board approval was obtained from each of the 11 participating sites. Descriptions of the inclusion/exclusion criteria of ONDRI participants were previously

reported (Farhan et al., 2017; Sunderland et al., 2022). All participants provided informed written consent. Clinical diagnoses and demographic data were obtained during participant screening and baseline assessment. When possible, participants provided clinical longitudinal follow-up assessment yearly, for up to 3 years (Farhan et al., 2017; Sunderland et al., 2022).

2.2 | Next-generation targeted sequencing

Of the 520 enrolled participants, 519 participants had a blood sample collected, from which genomic DNA was extracted. DNA was also obtained from 189 cognitively normal elderly controls from the GenADA study (Li et al., 2008).

All ONDRI participant and control DNA samples were subjected to targeted NGS using the ONDRISeq neurodegenerative disease gene panel, as previously described (Dilliott et al., 2018). DNA samples were pooled and paired-end NGS was performed using the MiSeq Personal Genome Sequencer (Illumina, San Diego, CA, United States) and MiSeq Reagent Kit v3. Raw sequencing data FASTQ files were imported into CLC Bio Genomics Workbench v10 (CLC Bio, Aarhus, Denmark) to perform preprocessing and variant annotation, which produced a variant calling format (VCF) file and binary alignment map (BAM) file for each participant. Read mapping was performed using the human reference genome GRCh37/hg19.

2.3 | CNV detection

The CNV Caller tool, an application within VarSeq® (v1.4.3; Golden Helix, Bozeman, MT), was used to detect CNVs from ONDRISeq-generated data. The CNV Caller tool employs a normalized depth of coverage algorithm, such that increase in sample coverage in comparison to a set of reference samples suggest a gain of copy number, and decrease in coverage suggests a loss of copy number.

The ONDRISeq browser extensible data (BED) file was imported into VarSeq®, as well as the VCF and BAM files of the 189 control samples from which the algorithm selected 48 to use as a reference set with the lowest percent difference in coverage data compared to each ONDRI sample. The algorithm excluded control samples in the reference set with >20% difference in coverage compared with the samples of interest. The matched reference sets also corrected for GC-content bias and regions exhibiting inaccurate mapping. By comparing to the reference set, the CNV Caller tool was used to identify CNVs across the 519 ONDRI participants. A DOC ratio and z-score were

computed for each target region covered by the NGS panel within each ONDRI sample. The DOC ratios measured the normalized DOC of the sample of interest compared to the normalized mean DOC of the reference set, whereas z-scores measured the number of standard deviations each target region's DOC was from the normalized mean DOC of the reference set. Additionally, the CNV Caller tool examined single nucleotide polymorphism (SNP) heterozygosity by examining variant allele frequencies across target regions to provide evidence for suspected CNVs, as previously described (Iacocca et al., 2017). The CNV Caller tool assigned each suspected CNV an average DOC ratio, average z-score, and a p-value. CNVs detected using ONDRISeq data are referred to as "potential CNVs".

2.4 | Deletion confirmation using breakpoint analysis

To verify the presence of a partial _OPTN_ (OMIM: 602432; NG_012876.1) gene deletion, primers were designed to flank regions surrounding putative deletion breakpoints and used for PCR amplification of the mutant allele. The Expand 20 kb^plus PCR system was used for DNA amplification (Roche, Basel, Switzerland). Forward (F3) and reverse (R1) primers flanking the deletion junction were: F3 5′-GTGACTCCATCACTCTGAACCTCC and R1 5′-CGAGTCTTCCTTCACATACGTGCC. Gel electrophoresis of the PCR product provided a visual confirmation of the mutant allele.

Once deletion breakpoints were identified, confirmation primers (P1:5′-TCCCTTGACATTTGCAGT GGAATC, P2: 5′-ACTGAGAGAACAGACAAGGTCAAC, P4: 5′-GGTCACTTAGGGAACAAGATAGTC) spanning proximal and distal breakpoints were designed for PCR and Sanger sequencing to verify the deletion breakpoint sequences for the wild type and mutant alleles. Thirty seconds of extension time for PCR cycles were used to achieve amplification of the normal allele using primer pair P1 and P2, while primer pair P1 and P4 amplified the mutant allele. Electropherograms were analyzed using the Applied Biosystems SeqScape Software (v2.6, Thermo Fischer Scientific, Waltham, MA, USA) with the reference sequence obtained from the National Center for Biotechnology Information GenBank database (https://www.ncbi.nlm.nih.gov/genbank/).

2.5 | Duplication confirmation using whole-exome sequencing

To validate presence of potential duplications, six samples, each with at least one potential duplication, were selected for whole-exome sequencing (WES). DNA samples were sent to the McGill University and Genome Quebec Innovation Centre (MUGQIC) for WES using the HiSeq 4000 instrument (Illumina) and Roche Nimblegen chemistry (Roche, Basel, Switzerland). FASTQ files were again imported into CLC Bio to perform preprocessing and variant annotation to produce a VCF file and BAM file for each participant.

VCF and BAM files of the six ONDRI participants and the BED file that defined the Roche Nimblegen chemistry target regions were imported into VarSeq®, along with VCF and BAM files from WES of eight reference samples obtained from cognitively normal individuals diagnosed with atrial fibrillation and sequenced on the same HiSeq 4000 run at the MUGQIC. Five reference samples were selected by the algorithm based on similarity of the normalized coverage to the samples of interest, as described above. Again, the CNV Caller tool applied a DOC approach and computed a DOC ratio, z-score, and p-value for each detected CNV.

3 | RESULTS

3.1 | Study participants and ONDRISeq CNV analysis

Using the VarSeq® CNV Caller tool, at least one potential CNV was detected in 44 of the total 519 ONDRI participants screened (8.5%; Table 1). A total of 47 potential CNVs were detected among the 44 participants, including 37 duplications and 10 heterozygous deletions (Appendix Table A1). The CNVs ranged in size from 150 to 74,407 bp.

3.2 | Deletion confirmation using breakpoint analysis

Of the 10 potential heterozygous deletions identified, one was chosen for breakpoint analysis based on our high confidence in the variant call, as determined by the metrics produced by the CNV Caller algorithm (DOC ratio = 0.487; z-score = −6.851; p-value = 1.10E−12). Specifically, the chosen heterozygous deletion encompassed exon 5 of _OPTN_. Sanger sequencing across the CNV breakpoints confirmed the presence of a 4969 bp deletion in _OPTN_ that encompassed all of exon five, with proximal and distal breakpoints at chr10: 13,152,598 and chr10: 13,157,566, respectively (Figure 1). The deletion was carried by subject 1, who was diagnosed with AD (Table 2). Although a deletion with these exact breakpoints has not been previously

TABLE 1 Demographics and CNV carrier status of the total ONDRI cohort

Cohort	Samples	Mean age (years ± SD)	Male: Female	Samples carrying potential CNVs (%)	Samples carrying validated CNVs (%)
ONDRI	519	68.6 ± 7.6	341:172	44 (8.5)	4 (0.8)
AD	41	71.8 ± 8.0	24:17	4 (9.8)	1 (2.4)
ALS	40	62.0 ± 8.7	24:16	4 (10.0)	1 (2.5)
CVD	161	69.2 ± 7.4	109:50	22 (13.7)	1 (0.6)
FTD	53	67.8 ± 7.1	34:19	1 (1.9)	0
MCI	85	70.6 ± 8.3	45:40	6 (7.1)	1 (1.2)
PD	139	67.8 ± 6.4	106:30	7 (5.0)	0

Abbreviations: AD, Alzheimer's disease; ALS, amyotrophic lateral sclerosis; CNV, copy number variant; CVD, cerebrovascular disease; FTD, frontotemporal dementia; MCI, mild cognitive impairment; ONDRI, Ontario neurodegenerative disease research initiative; PD, Parkinson's disease; SD, standard deviation.
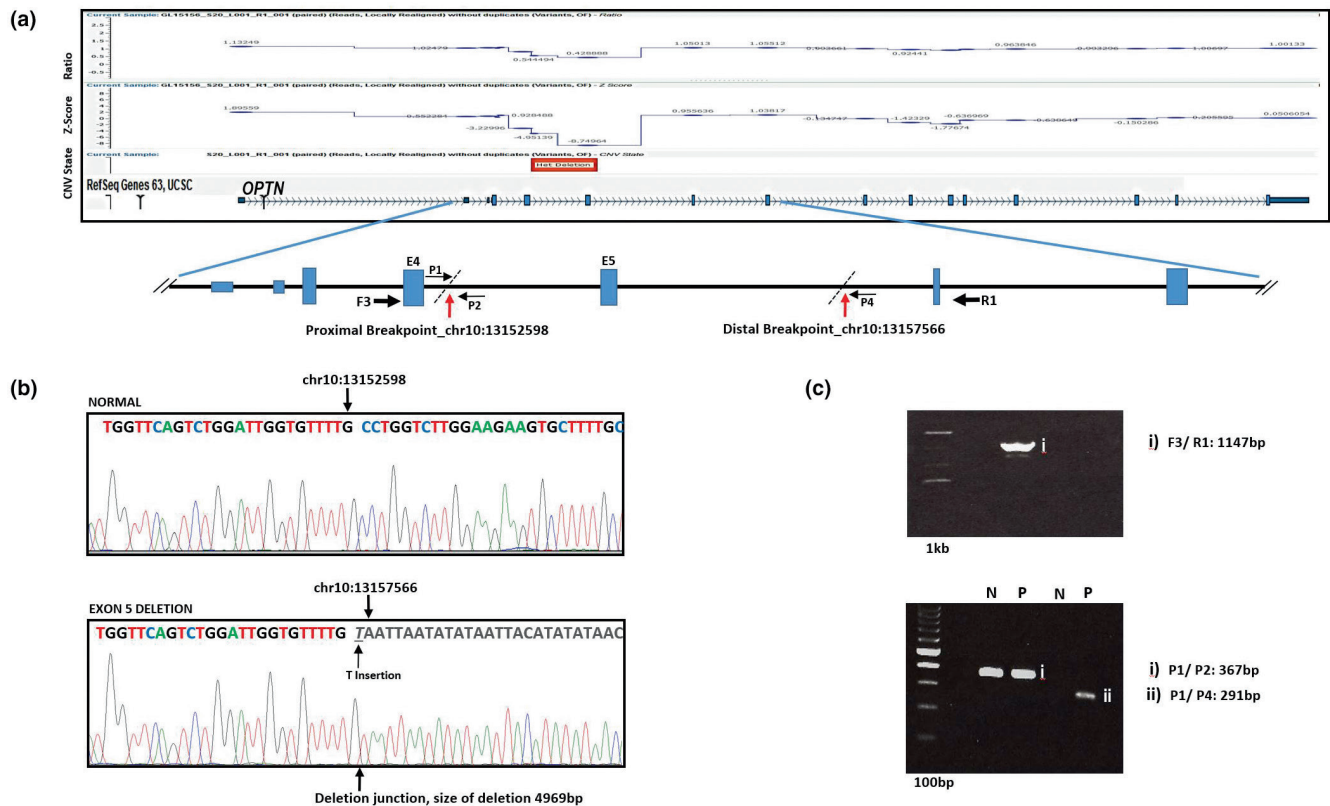

FIGURE 1 Validation of single-exon deletion in *OPTN* of subject 1 with Alzheimer's disease. (a) Screen capture of ONDRISeq-generated data from subject 1 processed by the VarSeq® v1.4.3 CNV caller tool identified a potential heterozygous deletion, as indicated by a drop in DOC ratio. The bottom section shows the *OPTN* gene and location of primers used to confirm and sequence across the breakpoint. (b) Sanger sequencing results for the deletion junction. Results from a cognitively normal control are presented on the top, with results from subject 1 on the bottom. Internal sequence missing in the deleted allele is written in gray. (c) Gel electrophoresis of PCR products across the deletion breakpoint. The top gel shows amplification products generated using F3 and R1. The normal sequence distance between primer pair F3 and R1 generated a product size of 6116 bp; however, PCR amplification of subject 1's genomic DNA using F3 and R1 generated a product size of 1147 bp, suggesting a 4969 bp deletion. The bottom gel contains amplification products generated using primer pairs (i) P1, located in the proximal side of the suspected breakpoint, and P2, located within the deleted fragment, as well as (ii) P1 and P4, located on the distal side of the suspected breakpoint. Both the normal control (N) and proband (P) demonstrate amplification (367 bp) for the proximal primer pair. Amplification (291 bp) with P1 and P4 is seen in the proband, but not the normal control. For individuals without the deletion, the span between P1 and P4 would be too large to amplify understandard conditions; thus, if amplification occurred, it confirms the presence of a large deletion between the primer pair.

reported within DECIPHER (Firth et al., 2009) (v11.7) or gnomAD (Karczewski et al., 2020) (v2.1 non-neuro; https://gnomad.broadinstitute.org/), a similar deletion was identified in one East-Asian individual within the gnomAD cohort (MAF = 4.14E−4) with breakpoints of chr10: 13,152,822 and chr10: 13,157,646.

The remaining nine heterozygous deletions did not undergo breakpoint analysis due to lower confidence in the reliability of the CNV calls, as determined by the CNV's individual metrics and/or relatively short span.

3.3 | Duplication confirmation using whole-exome sequencing

Of the participants harboring the 37 potential duplications, six were of relatively high confidence, as determined by their DOC ratios, z-scores, and p-values. Following WES of the samples and subsequent analysis with the CNV Caller tool, we identified and validated three participants as carriers of large-scale duplications (Table 2).

Subject 2 was diagnosed with ALS and harbored a duplication spanning 2731 bp that encompassed the first five exons of *PARK7* (chr1:8,021,464–8,031,243; OMIM: 602533; NG_008271.1). The duplication was not identified in DECIPHER or gnomAD, but a similar duplication had been previously reported as a variant of uncertain significance for PD within ClinVar. Subject 3 was diagnosed with CVD and harbored a duplication encompassing the entirety of *ABCC6* (OMIM: 603234; NG_007558.3), which was detected using the ONDRISeq analysis; however, WES revealed the duplication also encompassed 42 other neighboring genes, including 15 protein-coding genes, 12 pseudogenes, 11 microRNA encoding genes, and four non-coding RNA genes (Appendix Table A2). In total, the duplication spanned over 3 Mb (chr16:15,185,138–18,418,365). Although the specific CNV we identified has not been previously reported, a full-gene duplication of *ABCC6* was reported in gnomAD. The duplication was identified in 18 gnomAD samples (allele frequency = 1.08E−3); however, all individuals were under the age of 60 years, and it remains unclear whether any of the individuals presented with features of CVD. The duplication was also not reported in DECIPHER; however, three individuals were found to carry marginally smaller duplications of just under 3 Mb, and one individual carried a similarly sized duplication that was upstream of the one we identified, all of which encompassed *ABCC6*. Interestingly, all four of these duplications had been classified as likely pathogenic for phenotypes involving global developmental delay. Finally, subject 4 was diagnosed with MCI and harbored a duplication of exons 7–11 of *SAMHD1* (chr20:35,539,371–35,548,172; OMIM: 606754; NG_017059.1), which was not previously reported in DECIPHER or gnomAD.

One of the six samples sent for WES exhibited unmappable and incorrectly mapped reads failing to pass the quality control standards of the CNV Caller tool algorithm. Validation of the duplication carried by this individual remains inconclusive.

Importantly, none of the confirmed CNVs were identified in any of the 189 cognitively normal elderly control samples. The four confirmed CNV carriers did not harbor additional pathogenic single nucleotide variants in the neurodegenerative disease-associated genes encompassed by ONDRISeq relevant to their diagnoses. Clinical case information of subjects 1–4 is presented in Table 2.

4 | DISCUSSION

Of the 519 individuals diagnosed with neurodegenerative disease and/or CVD enrolled in the ONDRI study, we identified 44 (8.5%) with potential CNVs in the 80 neurodegenerative disease-associated genes covered by the ONDRISeq panel, of which four CNVs, each in a different participant (0.8%), were validated. Breakpoint analysis confirmed the presence of a heterozygous deletion in *OPTN* harbored by an individual with AD, and WES confirmed the presence of duplications in *PARK7*, *ABCC6*, and *SAMHD1*, in individuals diagnosed with AD, CVD, and MCI, respectively. To our knowledge, all four CNVs were novel with respect to each carrier's diagnosis.

Among the validated CNVs, the heterozygous deletion of exon 5 in *OPTN* was the only CNV we confirmed using Sanger-based breakpoint analysis and was identified in a participant with AD. *OPTN* encodes optineurin, and pathogenic variants in the gene are associated with both autosomal dominant, adult-onset glaucoma, and ALS (Maruyama et al., 2010; Schilter et al., 2015). Interestingly, two studies have reported similar heterozygous deletions of exon 5 of *OPTN* in Japanese ALS cohorts (Iida et al., 2012; Maruyama et al., 2010). It has been hypothesized that there may be a relationship between *OPTN*, glaucoma, and AD, due to the high rate of co-morbidity between glaucoma and AD, as well as the observation of optineurin in neurofibrillary tau tangles—a hallmark of AD pathology (Liu & Tian, 2011). Yet AD patients with pathogenic *OPTN* variants have not been previously observed. Herein, subject 1 exhibited gradual cognitive decline according to Montreal Cognitive Assessment (MoCA) scoring, but did not demonstrate ALS-associated motor symptoms. Although no glaucoma diagnosis was documented, the participant did report vision loss and presented with cataracts at baseline assessment. It remains unclear whether the observation of the partial *OPTN* heterozygous deletion may suggest a novel relationship between the gene and AD or whether the variant may be contributing to the participant's ocular phenotypes, and further functional analyses are required.

We confirmed the presence of a duplication spanning exons 1–5 of *PARK7* in a participant diagnosed with ALS using WES, although breakpoints could not be determined. *PARK7*, otherwise referred to as *DJ-1*, encodes a

TABLE 2 Demographics and clinical data of the six ONDRI participants identified to have validated CNVs in the 80 neurodegenerative disease genes covered by the ONDRISeq panel

	Subject 1	Subject 2	Subject 3	Subject 4
Diagnosis	AD	ALS	CVD	MCI
Age (years)	75	57	71	72
ASO (years)	73	55	70	57
Sex	Male	Male	Male	Female
MoCA	19	26	29	23
Other relevant clinical information	Exhibited cataracts in both eyes. No evidence of motor impairment reported	Exhibited slight kinetic tremor; fasciculation of the torso, arms, and legs; brisk deep tendon reflex; and diffuse denervation of the lower motor neurons. Other than a slight tremor, no signs of parkinsonism were reported (H&Y = 0)	Experienced right-sided, anterior, large-artery atherosclerosis. Exhibited coronary artery disease, hypertension, and high cholesterol and previously had undergone coronary artery bypass graft surgery	Although symptom onset was 15 years ago, the subject has not progressed to AD. No history of significant CVD.
Relevant family history	N/A	N/A	Both biological parents and one biological sibling exhibited heart disease	N/A
CNV Identified	Heterozygous deletion of exon 5 in *OPTN*	Duplication of exons 1–5 in *PARK7*	Duplication of all exons of *ABCC6* as well as 42 other genes[a]	Duplication of exons 7–11 in *SAMHD1*
Span (bp)	4,969[b]	9810	3,233,228[c]	8802
Genomic region	chr10:13,152,598–13,157,566	~chr1:8,021,464–8,031,273	~chr16:15,185,138–18,418,365[c]	~chr20:35,539,371–35,548,172
GenBank reference	NG_012876.1	NG_008271.1	NG_007558.3	NG_017059.1
Validation	Breakpoint	WES	WES	WES
DOC ratio	−6.851	6.816	6.371	5.872
z-score	0.487	1.455	1.482	1.474
p-value	1.10E−12	0	0	1.70E−21

Notes: Clinical data were obtained during participant screening and baseline visits of the ONDRI study. Span, genomic region, DOC ratio, z-score, and *p*-value were were based on the ONDRISeq CNV analysis, unless otherwise indicated. Genomic regions are in reference to GRCh37/hg19.

Abbreviations: AD, Alzheimer's disease; ALS, amyotrophic lateral sclerosis; ASO, age of symptom onset; bp, base pairs; chr, chromosome; CNV, copy number variant; CVD, cerebrovascular disease; DOC, depth of coverage; H&Y, Hoehn and Yahr; MCI, mild cognitive impairment; MoCA, Montreal Cognitive Assessment; N/A, not applicable; WES, whole-exome sequencing.

[a] The other genes encompassed by the duplication are outlined in Appendix Table A2.

[b] The span presented is based on the breakpoint analysis of the heterozygous deletion.

[c] The span and genomic region presented is based on the WES analysis of the duplication.

conserved protein belonging to the peptidase C56 family and is thought to inhibit aggregation of α-synuclein—a hallmark of PD pathology—as well as protect neurons against oxidative stress and cell death (Lev et al., 2006). Mutations in *PARK7* cause autosomal recessive early-onset PD (EOPD)(Bonifati et al., 2003). Previously, duplications of the first five exons of *PARK7* have been associated with EOPD (Macedo et al., 2009), and small-scale variants in the gene have been associated with autosomal recessive Guamanian ALS/EOPD, which presents with heterogeneous symptoms including muscular atrophy, cognitive decline, and tremor or rigidity (Annesi et al., 2005; Hanagasi et al., 2016). However, no previous associations of *PARK7* duplications have been reported in patients diagnosed with ALS alone. While the participant presented herein did report slight tremors at baseline, which progressed to moderate tremors upon 3-year follow-up based on the MDS Unified Parkinson's Disease Rating Scale (MDS-UPDRS), they had no further clinical signs of parkinsonism and had a Hoehn and Yahr score of zero both at baseline and follow-up. Yet the involvement of *PARK7* in ALS cannot be ruled out, as the DJ-1 protein is involved in sensing oxidative stress (Lev et al., 2006), and *PARK7* variants may therefore increase risk of oxidative stress, which is implicated as a major component in ALS pathologic mechanisms (Barber & Shaw, 2010). Our study is the first to report a case in which a structural variant affecting *PARK7* may have a role in a non-parkinsonism condition; however, functional analyses are required for further investigation of this relationship.

Furthermore, a CVD participant presented with a full-gene duplication of *ABCC6*. Confirmation of the duplication using WES found the CNV spanned over 3 Mb encompassing 43 total genes. Pathogenic variants within *ABCC6*, including CNVs, cause pseudoxanthoma elasticum (PXE), a rare autosomal recessive disorder characterized by elastic tissue fragmentation and arterial calcification (Bergen et al., 2000; Kringen et al., 2015; Ringpfeil et al., 2000). It is not uncommon for PXE patients to present with cerebral artery calcification, and studies have shown that ischemic CVD is highly prevalent in patients with PXE (Kauw et al., 2017; Pavlovic et al., 2005). Here, the participant harboring the *ABCC6* duplication presented with a history of conditions characteristic of PXE, including hypertension, atherosclerosis, stroke, mood disorders, and ocular features such as cataracts. The participant also reported that all immediate family members had a history of heart disease, albeit segregation analysis of the duplication was not possible. To our knowledge, this is the first reporting of a large-scale duplication involving *ABCC6* in an individual with CVD. Although this CNV spanned 42 other genes (Appendix Table A2), there is currently no evidence suggesting that

structural variation of these other genes contributed to the participant's disease presentation.

Finally, we identified and validated a duplication of exons 7–11 in *SAMHD1* in a subject with MCI. No reports of neurodegenerative symptoms have been made in patients demonstrating similar CNVs previously. Further evaluation of this CNV will be needed to gain a better understanding of its contribution to neurodegeneration, specifically cognitive impairment.

Although we have validated the presence of four CNVs across the ONDRI participants, we confirmed the exact breakpoints of only one, namely the deletion of exon 5 in *OPTN*. Identification of CNVs using NGS is limited to only determine which NGS probes are affected by the structural variant, thereby requiring further analysis to determine CNV breakpoints. However, breakpoint analysis remains challenging for duplications, as it is unclear whether the duplicated sequence will appear in tandem with the original sequence, or will be inserted unpredictably into a distal region of the genome. Therefore, we were unable to determine the exact location of the identified duplications and whether they may be interrupting other important genomic sequences that could contribute to the neurodegenerative phenotypes. Further, 43 CNVs identified using DOC analysis of ONDRISeq data remain unvalidated with average DOC ratios, average z-scores, and p-values of unknown confidence. Confirmation of these CNVs using alternative methods will be required. Despite these limitations, DOC CNV detection with targeted NGS continues to produce comprehensive, high-quality data, while remaining more time- and cost-effective than the "gold-standard" Sanger sequencing or multiplex ligation-dependent probe amplification approaches (Iacocca et al., 2017).

5 | CONCLUSION

In summary, we were able to identify and validate potentially pathogenic, novel CNVs in four individuals who were diagnosed with neurodegenerative disease or CVD. Further, we present an additional 43 potential CNVs that will be candidates for future replication studies. Although functional analyses are still required to determine how the CNVs may contribute to pathologic mechanisms of disease, the results highlight the need for further investigation into structural variants and their impact on neurodegenerative and cerebrovascular phenotypes. The CNVs may account for a portion of the missing heritability observed across the individual diagnoses. Assessing the full spectrum of potential variants that can contribute to the disease states is imperative for a complete understanding of the genetic etiology of these highly prevalent and progressive conditions, which, in due course, will contribute

to more accurate genetic diagnostic screening and therapeutic targeting.

The ONDRI Investigators:

Name	ONDRI Affiliation
Sabrina Adamo	Neuroimaging
Stephen Arnott	Neuroinformatics
Rob Bartha	Neuroimaging Lead
Derek Beaton	Neuroinformatics
Courtney Berezuk	Neuroimaging
Alanna Black	Gait and Balance
Alisia Bonnick	VCI Scholar
David Breen	PD Scholar
Don Brien	Eye Tracking
Susan Bronskill	Executive Member
Dennis Bulman	Genomics
Leanne Casaubon	VCI Clinician
Ying Chen	Eye Tracking
Marvin Chum	Recruitment: ALS
Brian Coe	Eye Tracking
Ben Cornish	Gait and Balance
Sherif Defrawy	SD-OCT
Jane Lawrence Dewar	Clinical
Roger A. Dixon	Pathology and Biomarkers
Sali M.K. Farhan	Genomics
Frederico Faria	Gait and Balance
Julia Fraser	Gait and Balance
Mahdi Ghani	Genomics
Barry Greenberg	SD-OCT Lead
Hassan Haddad	Neuroimaging
Wendy Hatch	SD-OCT
Melissa Holmes	Neuroimaging
Chris Hudson	SD-OCT Lead
Peter Kleinstiver	Program Manager
Donna Kwan	Neuropsychology and Clinical
Elena Leontieva	SD-OCT

Name	ONDRI Affiliation
Brian Levine	Neurpsychology
Wendy Lou	Neuroinformatics
Efrem Mandelcorn	SD-OCT
Ed Margolin	SD-OCT
Connie Marras	Recruitment: PD
Bill McIlroy	Gait and Balance Lead
Paula McLaughlin	Neuropsychology
Manuel Montero Odasso	Gait and Balance Lead
Doug Munoz	Eye Tracking Lead
David Munoz	Pathology and Biomarkers Lead
Nuwan Nanayakkara	Neuroimaging
JB Orange	Neuropsychology
Miracle Ozzoude	Neuroimaging
Alicia Peltsch	Neuropsychology
Pradeep Raamana	Neuroinformatics
Joel Ramirez	Neuroimaging
Natalie Rashkovan	Program Manager
Angela Roberts	Neuropsychology
Yanina Sarquis Adamson	Gait and Balance
Christopher Scott	Neuroimaging
Michael Strong	Program Lead
Stephen Strothers	Neuroimaging/Neuroinformatics Lead
Sujeevini Sujanthan	Neuroinformatics
Kelly M. Sunderland	Neuroinformatics
Sean Symons	Neuroimaging Lead
Faryan Tayyari	SD-OCT
Athena Theyers	Neuroinformatics
Angela Troyer	Neuropsychology
Abiramy Uthirakumaran	Neuroinformatics
Karen Van Ooteghem	Gait and Balance
John Woulfe	Pathology and Biomarkers
Mojdeh Zamyadi	Neuroinformatics
Guangyong (GY) Zou	Neuroinformatics

CONFLICT OF INTERESTS

The authors declare the following financial interests/personal relationships, none of which are in conflict with the work presented herein: SEB has peer reviewed operating grants to the institution as PI/CoPI from CIHR, American Drug Discovery Foundation, National Institute of Health, Weston Brain Institute, Leducq Foundation, Brain Canada Platform grant, the Surgeon General, Kuwait Ministry of Health, Canadian Foundation for Innovation. SEB is the Site PI with funding to the institution from Hoffman La Roche, Biogen, and Eli Lilly. EF has received personal compensation for serving on a PSP scientific advisory board for Biogen, for serving as a section editor for NeuroImage Clinical, and for serving as a course director for the AAN annual meeting. CEF receives grant funding from Hoffman La Rcohe, Vielight Inc., Cortexyme, Brain Canada, NIH, CIHR—Canadian Consortium on Neurodegeneration in Aging, and the St. Michaels Hospital Foundation Heather and Eric Donnelly endowment outside of the submitted work. DG reports honorariums from Sunovion and Paladin Labs Inc. as well as clinical trials with CIHR, Genzyme Corporation/Sanofi Canada, and Eli Lilly and Company. SK has received research support from Brain and Behavior Foundation, National institute on Aging, BrightFocus Foundation, Brain Canada, Canadian Institute of Health Research, Canadian Consortium on Neurodegeneration in Aging, Centre for Aging and Brain Health Innovation, Centre for Addiction and Mental Health, and University of Toronto. AEL reports consultancy support from Abbvie, Acorda, AFFiRis, Biogen, Denali, Janssen, Intracellular, Kallyope, Lundbeck, Paladin, Retrophin, Roche, Sun Pharma, Theravance, and Corticobasal Degeneration Solutions; advisory board support form Jazz Pharma, PhotoPharmics, Sunovion; other honoraria from Sun Pharma, AbbVie, Sunovion, American Academy of Neurology and the International Parkinson and Movement Disorder Society; grants from Brain Canada, Canadian Institutes of Health Research, Corticobasal Degeneration Solutions, Edmond J Safra Philanthropic Foundation, Michael J. Fox Foundation, the Ontario Brain Institute, Parkinson Foundation, Parkinson Canada, and W. Garfield Weston Foundation and royalties from Elsevier, Saunders, Wiley-Blackwell, Johns Hopkins Press, and Cambridge University Press. MM receives salary support from the Department of Medicine at Sunnybrook Health Sciences Centre, the University of Toronto, and the Sunnybrook Research Institute, as well as advisory board support from Arkuda Therapeutics, Ionis

Pharmaceuticals, Alector and Wave Life Sciences; royalties from Henry Stewart Talks Ltd; grants paid to institution from Weston Brain Institute, Ontario Brain Institute, Washington University and Canadian Institutes of Health Research outside the submitted work; and clinical trials support from Roche and Alector. GS is the Editor-in-Chief of the World Stroke Academy for the World Stroke Organization and receives CME honorarium from Servier and Roche, as well as grants from Roche. MCT receives consultancy support from Biogen and Hoffman-La Roche and is a board member of Alzheimer's Society of Toronto.

ETHICAL APPROVAL
Study ethics approval was obtained from the Research Ethics Boards at Baycrest Centre for Geriatric Care (Toronto, Ontario, Canada); Centre for Addiction and Mental Health (Toronto, Ontario, Canada); Elizabeth Bruyère Hospital (Ottawa, Ontario, Canada); Hamilton General Hospital (Hamilton, Ontario, Canada); McMaster (Hamilton, Ontario, Canada); London Health Sciences Centre (London, Ontario, Canada); Parkwood Hospital (London, Ontario, Canada); St Michael's Hospital (Toronto, Ontario, Canada); Sunnybrook Health Sciences Centre (Toronto, Ontario, Canada); The Ottawa Hospital (Ottawa, Ontario, Canada); and University Health Network-Toronto Western Hospital (Toronto, Ontario, Canada). All participants provided written, informed consent in accordance with the Research Ethics Boards and regulatory requirements.

DATA AVAILABILITY STATEMENT
In accordance with the Ontario Neurodegenerative Disease Research Initiative (ONDRI) with the Ontario Brain Institute, all baseline data from ONDRI are available upon request at https://www.braincode.ca/. To gain access to the data, an account request must be made to help@brain-code.ca. Full details regarding data access can be found at https://www.braincode.ca/content/getting-started.

ORCID
Allison A. Dilliott https://orcid.org/0000-0003-3863-9304
Robert A. Hegele https://orcid.org/0000-0003-2861-5325

REFERENCES
Annesi, G., Savettieri, G., Pugliese, P., D'Amelio, M., Tarantino, P., Ragonese, P., La Bella, V., Piccoli, T., Civitelli, D., Annesi, F., & Fierro, B. (2005). DJ-1 mutations and parkinsonism-dementia-amyotrophic lateral sclerosis complex. *Annals of Neurology*, *58*(5), 803–807. https://doi.org/10.1002/ana.20666

Barber, S. C., & Shaw, P. J. (2010). Oxidative stress in ALS: Key role in motor neuron injury and therapeutic target. *Free Radical Biology & Medicine*, *48*(5), 629–641. https://doi.org/10.1016/j.freeradbiomed.2009.11.018

Bergen, A. A., Plomp, A. S., Schuurman, E. J., Terry, S., Breuning, M., Dauwerse, H., Swart, J., Kool, M., van Soest, S., Baas, F., &

ten Brink, J. B. (2000). Mutations in ABCC6 cause pseudoxanthoma elasticum. *Nature Genetics*, *25*(2), 228–231. https://doi.org/10.1038/76109

Bonifati, V., Rizzu, P., Squitieri, F., Krieger, E., Vanacore, N. A., Van Swieten, J. C., Brice, A., Van Duijn, C. M., Oostra, B., Meco, G., & Heutink, P. (2003). DJ-1(PARK7), a novel gene for autosomal recessive, early onset parkinsonism. *Neurological Sciences*, *24*(3), 159–160. https://doi.org/10.1007/s10072-003-0108-0

Cacace, R., Sleegers, K., & Van Broeckhoven, C. (2016). Molecular genetics of early-onset Alzheimer's disease revisited. *Alzheimers Dement*, *12*(6), 733–748. https://doi.org/10.1016/j.jalz.2016.01.012

Cuccaro, D., De Marco, E. V., Cittadella, R., & Cavallaro, S. (2017). Copy number variants in Alzheimer's disease. *Journal of Alzheimer's Disease*, *55*(1), 37–52. https://doi.org/10.3233/JAD-160469

Dilliott, A. A., Farhan, S. M. K., Ghani, M., Sato, C., Liang, E., Zhang, M., McIntyre, A. D., Cao, H., Racacho, L., Robinson, J. F., Strong, M. J., Masellis, M., Bulman, D. E., Rogaeva, E., Lang, A., Tartaglia, C., Finger, E., Zinman, L., Turnbull, J., ... Hegele, R. A. (2018). Targeted next-generation sequencing and bioinformatics pipeline to evaluate genetic determinants of constitutional disease. *J Vis Exp*, *134*, 57266. https://doi.org/10.3791/57266

Farhan, S. M. K., Bartha, R., Black, S. E., Corbett, D., Finger, E., Freedman, M., Greenberg, B., Grimes, D. A., Hegele, R. A., Hudson, C., Kleinstiver, P. W., Lang, A. E., Masellis, M., McIlroy, W. E., McLaughlin, P. M., Montero-Odasso, M., Munoz, D. G., Munoz, D. P., Strother, S., ... Strong, M. J. (2017). The Ontario neurodegenerative disease research initiative (ONDRI). *The Canadian Journal of Neurological Sciences*, *44*(2), 196–202. https://doi.org/10.1017/cjn.2016.415

Farhan, S. M. K., Dilliott, A. A., Ghani, G., Sato, C., Liang, E., Zhang, M., McIntyre, A. D., Cao, H., Racacho, L., Robinson, J. F., Strong, M. J., Masellis, M., St George-Hyslop, P., Bulman, D. E., Rogaeva, E., & Hegele, R. A. (2016). The ONDRISeq panel: Custom-designed next-generation sequencing of genes related to neurodegeneration. *NPJ Genom Med*, *16032*, 1–11.

Feuk, L., Carson, A. R., & Scherer, S. W. (2006). Structural variation in the human genome. *Nature Reviews. Genetics*, *7*(2), 85–97. https://doi.org/10.1038/nrg1767

Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D., Van Vooren, S., Moreau, Y., Pettett, R. M., & Carter, N. P. (2009). DECIPHER: Database of chromosomal imbalance and phenotype in humans using Ensembl resources. *American Journal of Human Genetics*, *84*(4), 524–533. https://doi.org/10.1016/j.ajhg.2009.03.010

Ghani, M., Pinto, D., Lee, J. H., Grinberg, Y., Sato, C., Moreno, D., Scherer, S. W., Mayeux, R., George Hyslop, P., & Rogaeva, E. (2012). Genome-wide survey of large rare copy number variants in Alzheimer's disease among Caribbean hispanics. *G3 (Bethesda)*, *2*(1), 71–78. https://doi.org/10.1534/g3.111.000869

Greaves, C. V., & Rohrer, J. D. (2019). An update on genetic frontotemporal dementia. *Journal of Neurology*, *266*(8), 2075–2086. https://doi.org/10.1007/s00415-019-09363-4

Hagenaars, S. P., Radakovic, R., Crockford, C., Fawns-Ritchie, C., International, F. T. D. G. C., Harris, S. E., Gale, C.R., & Deary, I. J. (2018). Genetic risk for neurodegenerative disorders, and its overlap with cognitive ability and physical function.

PLoS One, *13*(6), e0198187. https://doi.org/10.1371/journal.pone.0198187

Hanagasi, H. A., Giri, A., Kartal, E., Guven, G., Bilgic, B., Hauser, A. K., Emre, M., Heutink, P., Basak, N., Gasser, T., Simón-Sánchez, J., & Lohmann, E. (2016). A novel homozygous DJ1 mutation causes parkinsonism and ALS in a Turkish family. *Parkinsonism & Related Disorders*, *29*, 117–120. https://doi.org/10.1016/j.parkreldis.2016.03.001

Hooli, B. V., Kovacs-Vajna, Z. M., Mullin, K., Blumenthal, M. A., Mattheisen, M., Zhang, C., Lange, C., Mohapatra, G., Bertram, L., & Tanzi, R. E. (2014). Rare autosomal copy number variations in early-onset familial Alzheimer's disease. *Molecular Psychiatry*, *19*(6), 676–681. https://doi.org/10.1038/mp.2013.77

Iacocca, M. A., Wang, J., Dron, J. S., Robinson, J. F., McIntyre, A. D., Cao, H., & Hegele, R. A. (2017). Use of next-generation sequencing to detect LDLR gene copy number variation in familial hypercholesterolemia. *Journal of Lipid Research*, *58*(11), 2202–2209. https://doi.org/10.1194/jlr.D079301

Iida, A., Hosono, N., Sano, M., Kamei, T., Oshima, S., Tokuda, T., Nakajima, M., Kubo, M., Nakamura, Y., & Ikegawa, S. (2012). Novel deletion mutations of OPTN in amyotrophic lateral sclerosis in Japanese. *Neurobiology of Aging*, *33*(8), 1843 e1819-1824. https://doi.org/10.1016/j.neurobiolaging.2011.12.037

Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., Collins, R. L., Laricchia, K. M., Ganna, A., Birnbaum, D. P., Gauthier, L. D., Brand, H., Solomonson, M., Watts, N. A., Rhodes, D., Singer-Berk, M., England, E. M., Seaby, E. G., Kosmicki, J. A., ... MacArthur, D. G. (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*, *581*(7809), 434–443. https://doi.org/10.1038/s41586-020-2308-7

Karlsson, I. K., Escott-Price, V., Gatz, M., Hardy, J., Pedersen, N. L., Shoai, M., & Reynolds, C. A. (2022). Measuring heritable contributions to Alzheimer's disease: Polygenic risk score analysis with twins. *Brain Commun*, *4*(1), fcab308. https://doi.org/10.1093/braincomms/fcab308

Kauw, F., Kranenburg, G., Kappelle, L. J., Hendrikse, J., Koek, H. L., Visseren, F. L. J., Mali, W. P. T., de Jong, P. A., & Spiering, W. (2017). Cerebral disease in a nationwide Dutch pseudoxanthoma elasticum cohort with a systematic review of the literature. *Journal of the Neurological Sciences*, *373*, 167–172. https://doi.org/10.1016/j.jns.2016.12.053

Keller, M. F., Saad, M., Bras, J., Bettella, F., Nicolaou, N., Simon-Sanchez, J., Mittag, F., Büchel, F., Sharma, M., Gibbs, J. R., Schulte, C., Moskvina, V., Durr, A., Holmans, P., Kilarski, L. L., Guerreiro, R., Hernandez, D. G., Brice, A., Ylikotila, P., ... Wellcome Trust Case Control C. (2012). Using genome-wide complex trait analysis to quantify 'missing heritability' in Parkinson's disease. *Human Molecular Genetics*, *21*(22), 4996–5009. https://doi.org/10.1093/hmg/dds335

Kringen, M. K., Stormo, C., Berg, J. P., Terry, S. F., Vocke, C. M., Rizvi, S., Hendig, D., & Piehler, A. P. (2015). Copy number variation in the ATP-binding cassette transporter ABCC6 gene and ABCC6 pseudogenes in patients with pseudoxanthoma elasticum. *Molecular Genetics & Genomic Medicine*, *3*(3), 233–237. https://doi.org/10.1002/mgg3.137

Lev, N., Roncevic, D., Ickowicz, D., Melamed, E., & Offen, D. (2006). Role of DJ-1 in Parkinson's disease. *Journal of Molecular Neuroscience*, *29*(3), 215–225. https://doi.org/10.1385/jmn:29:3:215

Li, H., Wetten, S., Li, L., St Jean, P. L., Upmanyu, R., Surh, L., Hosford, D., Barnes, M. R., Briley, J. D., Borrie, M., Coletta, N., Delisle, R., Dhalla, D., Ehm, M. G., Feldman, H. H., Fornazzari, L., Gauthier, S., Goodgame, N., Guzman, D., ... Roses, A. D. (2008). Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease. *Archives of Neurology*, *65*(1), 45–53. https://doi.org/10.1001/archneurol.2007.3

Liu, Y. H., & Tian, T. (2011). Hypothesis of optineurin as a new common risk factor in normal-tension glaucoma and Alzheimer's disease. *Medical Hypotheses*, *77*(4), 591–592. https://doi.org/10.1016/j.mehy.2011.06.040

Macedo, M. G., Verbaan, D., Fang, Y., van Rooden, S. M., Visser, M., Anar, B., Uras, A., Groen, J. L., Rizzu, P., van Hilten, J. J., & Heutink, P. (2009). Genotypic and phenotypic characteristics of Dutch patients with early onset Parkinson's disease. *Movement Disorders*, *24*(2), 196–203. https://doi.org/10.1002/mds.22287

Maruyama, H., Morino, H., Ito, H., Izumi, Y., Kato, H., Watanabe, Y., Kinoshita, Y., Kamada, M., Nodera, H., Suzuki, H., Komure, O., Matsuura, S., Kobatake, K., Morimoto, N., Abe, K., Suzuki, N., Aoki, M., Kawata, A., Hirai, T., ... Kawakami, H. (2010). Mutations of optineurin in amyotrophic lateral sclerosis. *Nature*, *465*(7295), 223–226. https://doi.org/10.1038/nature08971

Mejzini, R., Flynn, L. L., Pitout, I. L., Fletcher, S., Wilton, S. D., & Akkari, P. A. (2019). ALS genetics, mechanisms, and therapeutics: Where are we now? *Frontiers in Neuroscience*, *13*, 1310. https://doi.org/10.3389/fnins.2019.01310

Morello, G., Guarnaccia, M., Spampinato, A. G., La Cognata, V., D'Agata, V., & Cavallaro, S. (2018). Copy number variations in amyotrophic lateral sclerosis: Piecing the mosaic tiles together through a systems biology approach. *Molecular Neurobiology*, *55*(2), 1299–1322. https://doi.org/10.1007/s12035-017-0393-x

Nuytemans, K., Theuns, J., Cruts, M., & Van Broeckhoven, C. (2010). Genetic etiology of Parkinson disease associated with mutations in the SNCA, PARK2, PINK1, PARK7, and LRRK2 genes: A mutation update. *Human Mutation*, *31*(7), 763–780. https://doi.org/10.1002/humu.21277

Pavlovic, A. M., Zidverc-Trajkovic, J., Milovic, M. M., Pavlovic, D. M., Jovanovic, Z., Mijajlovic, M., Petrovic, M., Kostic, V. S., & Sternic, N. (2005). Cerebral small vessel disease in pseudoxanthoma elasticum: Three cases. *The Canadian Journal of Neurological Sciences*, *32*(1), 115–118. https://doi.org/10.1017/s0317167100016991

Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., Fiegler, H., Shapero, M. H., Carson, A. R., Chen, W., Cho, E. K., Dallaire, S., Freeman, J. L., González, J. R., Gratacòs, M., Huang, J., Kalaitzopoulos, D., Komura, D., MacDonald, J. R., ... Hurles, M. E. (2006). Global variation in copy number in the human genome. *Nature*, *444*(7118), 444–454. https://doi.org/10.1038/nature05329

Ringpfeil, F., Lebwohl, M. G., Christiano, A. M., & Uitto, J. (2000). Pseudoxanthoma elasticum: Mutations in the MRP6 gene encoding a transmembrane ATP-binding cassette (ABC) transporter. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(11), 6001–6006. https://doi.org/10.1073/pnas.100041297

Schilter, K. F., Reis, L. M., Sorokina, E. A., & Semina, E. V. (2015). Identification of an Alu-repeat-mediated deletion of OPTN

upstream region in a patient with a complex ocular phenotype. *Molecular Genetics & Genomic Medicine*, *3*(6), 490–499. https://doi.org/10.1002/mgg3.159

Stranger, B. E., Forrest, M. S., Dunning, M., Ingle, C. E., Beazley, C., Thorne, N., Redon, R., Bird, C. P., de Grassi, A., Lee, C., Tyler-Smith, C., Carter, N., Scherer, S. W., Tavaré, S., Deloukas, P., Hurles, M. E., & Dermitzakis, E. T. (2007). Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science*, *315*(5813), 848–853. https://doi.org/10.1126/science.1136678

Sunderland, K. M., Beaton, D., Arnott, S. R., Kleinstiver, P., Kwan, D., Lawrence-Dewar, J. M., Ramirez, J., Tan, B., Bartha, R., Black, S. E., Borrie, M., Brien, D., Casaubon, L. K., Coe, B. C., Cornish, B., Dilliott, A. A., Dowlatshahi, D., Finger, E., Fischer, C., ... Binns, M. A. (2022). Characteristics of the Ontario Neurodegenerative Disease Research Initiative cohort. *Alzheimer's & Dementia*, *1*–*18*, https://doi.org/10.1002/alz.12632

APPENDIX A

Targeted copy number variant identification across the neurodegenerative disease spectrum

Tables A1 and A2

TABLE A1 CNVs detected by the VarSeq® v1.4.3 CNV caller tool in 519 ONDRI participants in 80 neurodegenerative disease genes covered by the ONDRISeq panel

Subject	Phenotype	CNV state	Gene	Exon (s)	Span (bp)	Z-score	Ratio	p-value	Validation	gnomAD[c] (N)	DECIPHER[c] (N)
1	AD	Het Deletion	OPTN	5	2357	−6.851	0.487	1.10E−12	Yes (BA)	1	0
2	ALS	Duplicate	PARK7	1–5	9810	6.816	1.455	0	Yes (WES)	0	0
3	CVD	Duplicate	ABCC6	1–31	74,407	6.371	1.482	0	Yes (WES)	18	4
4	MCI	Duplicate	SAMHD1	7–11	8802	5.872	1.474	1.70E−21	Yes (WES)	0	0
5	CVD	Duplicate	ABCA7	16	276	9.961	3.221	3.14E−13	N/A	0	0
6	CVD	Duplicate	ATP13A2	16	707	7.175	1.430	2.02E−08	N/A	0	0
7	CVD	Duplicate	ABCA7	4–5	611	4.156	1.357	3.51E−07	No (WES)	0	0
7	CVD	Het Deletion	LRRK2	9	170	−3.990	0.636	4.09E−05	N/A[a]	0	0
7	CVD	Duplicate	PRPH	7–8	453	3.991	1.338	6.34E−05	No (WES)	0	0
8	AD	Duplicate	ABCA7	15–16	630	4.454	1.845	4.63E−07	N/A	0	0
9	CVD	Het Deletion	TREM2	4	156	−3.753	0.678	7.30E−07	N/A	0	0
10	CVD	Duplicate	COL4A2	2	220	4.212	1.382	7.34E−07	N/A	0	0
11	PD	Duplicate	UNC13A	36	530	6.097	1.356	8.18E−07	N/A	0	0
12	CVD	Duplicate	SQSTM1	0	532	6.045	1.712	8.62E−07	N/A	0	0
13	PD	Duplicate	ABCA7	16–17	1627	4.163	1.748	9.65E−07	N/A	0	0
14	CVD	Duplicate	CLU	8	395	4.731	1.405	1.72E−06	N/A	0	0
15	ALS	Duplicate	DNMT1	35–36	678	6.623	1.319	2.14E−06	N/A	0	0
16	FTD	Het Deletion	TREM2	4	156	−3.980	0.653	3.77E−06	N/A	0	0
17	CVD	Duplicate	ATP13A2	28	700	5.397	1.493	4.39E−06	N/A	0	0
18	CVD	Duplicate	CENPV	1	972	5.323	1.441	5.14E−06	N/A	0	0
18	CVD	Duplicate	PSEN2	1	578	4.559	1.350	9.32E−05	N/A	0	0
19	CVD	Duplicate	NOTCH3	3–4	641	3.901	1.351	5.44E−06	N/A	0	0
20	PD	Duplicate	GCH1	5	621	6.382	1.322	8.90E−06	N/A[b]	0	0
21	PD	Duplicate	UCHL1	0–1	150	4.799	2.454	9.01E−06	N/A	0	0
22	PD	Duplicate	ALS2	34	2127	4.658	1.380	1.07E−05	N/A	0	0
23	MCI	Duplicate	TAF15	15	210	4.235	1.513	1.26E−05	N/A	0	0
24	CVD	Duplicate	NOTCH3	4–6	944	3.946	1.304	1.33E−05	No (WES)	0	0

TABLE A1 (Continued)

Subject	Phenotype	CNV state	Gene	Exon (s)	Span (bp)	Z-score	Ratio	p-value	Validation	gnomAD[c] (N)	DECIPHER[c] (N)
25	MCI	Het Deletion	GAK	20–23	623	−5.521	0.567	1.35E−05	N/A	12	0
26	CVD	Duplicate	ABCC6	3	265	4.652	1.786	1.40E−05	N/A	0	0
27	CVD	Duplicate	PNPLA6	1–3	375	3.722	1.596	1.48E−05	N/A	0	0
28	CVD	Duplicate	CD2AP	1	960	3.853	1.407	1.85E−05	N/A	0	0
29	AD	Duplicate	UNC13A	36	530	5.228	1.304	2.72E−05	N/A	0	0
30	ALS	Duplicate	COL4A2	7	3715	3.817	1.314	2.87E−05	N/A	0	0
31	CVD	Duplicate	SOD1	1–2	4765	3.541	2.019	3.13E−05	N/A	0	0
32	CVD	Duplicate	UNC13A	35	2216	4.415	1.374	3.21E−05	N/A	0	0
33	CVD	Het Deletion	NEFH	1	571	−5.899	0.564	3.22E−05	N/A	0	0
34	CVD	Duplicate	NOTCH3	21–24	1756	3.884	1.322	3.29E−05	N/A	0	0
35	ALS	Duplicate	UCHL1	7–8	2361	3.981	1.301	3.55E−05	N/A	0	0
36	CVD	Duplicate	NOTCH3	3–4	641	3.543	1.323	3.63E−05	N/A	0	0
37	MCI	Duplicate	TARDBP	1	622	5.506	1.426	4.24E−05	N/A	0	0
38	AD	Duplicate	ABCA7	13–16	1925	3.568	1.388	5.11E−05	N/A	0	0
38	CVD	Het Deletion	SIGMAR1	1	351	−3.757	0.538	5.76E−05	N/A	0	0
40	CVD	Duplicate	ATP13A2	1	700	5.030	1.460	6.29E−05	N/A	0	0
41	MCI	Duplicate	PNPLA6	5	232	5.132	1.421	7.97E−05	N/A	0	0
42	PD	Het Deletion	UNC13A	8	506	−4.214	0.598	8.88E−05	N/A	0	0
43	MCI	Het Deletion	ABCC6	3–4	629	−3.744	0.516	9.89E−05	N/A	0	0
44	PD	Het Deletion	ABCA7	8	222	−4.021	0.451	9.94E−05	N/A	0	0

Notes: For multi-exon CNVs, the reported ratio and Z-score values are averaged across each affected region. A response of "No" in respect to validation indicated that the WES did not identify the CNV that had been identified using the ONDRISeq panel.

Abbreviations: AD, Alzheimer's disease; ALS, amyotrophic lateral sclerosis; BA, breakpoint analysis; bp, base pairs; CNV, copy number variant; CVD, cerebrovascular disease; FTD, frontotemporal dementia; Het, heterozygous; MCI, mild cognitive impairment; N/A, not applicable; PD, Parkinson's disease; WES, whole-exome sequencing.

[a]The WES performed did not have probes adequately covering exon 9 of LRRK2.

[b]WES exhibited unmappable and incorrectly mapped reads, failing to pass the quality control standards of the CNV Caller tool algorithm.

[c]The CNV was considered observed in gnomAD (v2.1 non-neuro) or DECIPHER (v11.7) if the CNV in the database displayed similar breakpoints (i.e., covered the same exons of the affected gene).

TABLE A2 Genes encompassed by the >3 mb duplication harbored by subject 3

Gene type	Genes
MicroRNA encoding genes	*MIR1972-1; MIR6511B2; MIR3180-4; MIR6506; MIR484; MIR3179-2; MIR3670-2; MIR3180-2; MIR6511A2; MIR6770-2; MIR6511A3*
Noncoding RNA encoding genes	*LOC100505915; PKD1P6-NPIPP1; PKD1P1; LOC102723692*
Protein-coding genes	*PDXPC1; NTAN1; RRN3; NPIPA5; MPV17L; C16orf45; KIAA0430; NDE1; MYH11; FOPNL; ABCC1; ABCC6; NOMO3; NPIPA7; XYLT1; NPIPA8*
Pseudogenes	*LOC728138; NPIPP1; PKD1P6; RNU6-213P; RPL15P20; RPL17P40; PKD1P2; LOC100133127; LOC441750; LOC100421029; RPL7P47; LOC100133137*