The ONDRISeq panel: custom-designed next-generation sequencing of genes related to neurodegeneration

Sali MK Farhan1,2, Allison A Dilliott1,2, Mahdi Ghani3, Christine Sato3, Eric Liang3, Ming Zhang3, Adam D McIntyre1, Henian Cao1, Lemuel Racacho4,5, John F Robinson1, ONDRI Investigators, Michael J Strong1,6, Mario Maselli5, Peter St George-Hyslop3,8, Dennis E Bulman4,9, Ekaterina Rogavea3 and Robert A Hegele1,2

The Ontario Neurodegenerative Disease Research Initiative (ONDRI) is a multimodal, multi-year, prospective observational cohort study to characterise five diseases: (1) Alzheimer’s disease (AD) or amnestic single or multidomain mild cognitive impairment (aMCI) (AD/MCI); (2) amyotrophic lateral sclerosis (ALS); (3) frontotemporal dementia (FTD); (4) Parkinson’s disease (PD); and (5) vascular cognitive impairment (VCI). The ONDRI Genomics subgroup is investigating the genetic basis of neurodegeneration. We have developed a custom next-generation-sequencing-based panel, ONDRISeq that targets 80 genes known to be associated with neurodegeneration. We processed DNA collected from 216 individuals diagnosed with one of the five diseases, on ONDRISeq. All runs were executed on a MiSeq instrument and subjected to rigorous quality control assessments. We also independently validated a subset of the variant calls using NeuroX (a genome-wide array for neurodegenerative disorders), TaqMan allelic discrimination assay, or Sanger sequencing. ONDRISeq consistently generated high-quality genotyping calls and on average, 92% of targeted bases are covered by at least 30 reads. We also observed 100% concordance for the variants identified via ONDRISeq and validated by other genomic technologies. We were successful in detecting known as well as novel rare variants in 72.2% of cases although not all variants are disease-causing. Using ONDRISeq, we also found that the APOE E4 allele had a frequency of 0.167 in these samples. Our optimised workflow highlights next-generation sequencing as a robust tool in elucidating the genetic basis of neurodegenerative diseases by screening multiple candidate genes simultaneously.

INTRODUCTION

Dementia encompasses a heterogeneous group of neurodegenerative diseases characterised by a progressive decline in cognitive function, language deficiency, and in some cases, motor impairment and behavioural anomalies. Currently, dementia has a global prevalence of 47.5 million cases and an incidence of 7.7 million new cases annually.1–3 Although today there are no direct treatments available to alter the progressive disease course, early diagnosis has been one of the best predictors of disease outcome.3,4 Further understanding of the molecular basis of dementia can lead to earlier diagnosis and the eventual development of targeted and efficacious treatment modalities.

Our group is part of the Ontario Neurodegenerative Disease Research Initiative (ONDRI), a multimodal, multi-year, prospective observational cohort study designed to address the effect of small vessel disease in neurodegeneration. ONDRI is recruiting ~600 participants diagnosed with one of the following five diseases: (1) Alzheimer’s disease (AD) or amnestic single- or multidomain mild cognitive impairment (aMCI) (AD/MCI); (2) amyotrophic lateral sclerosis (ALS); (3) frontotemporal dementia (FTD); (4) Parkinson’s disease (PD); and (5) vascular cognitive impairment (VCI).

Genetics is an important risk factor for neurodegenerative disease. Approximately 5–10% of cases with neurodegenerative diseases are familial and can be attributed to several genes.5–7 However, it is likely we are underestimating the incidence of familial cases based on clinical ascertainment, as the death of presymptomatic individuals may be due to other medical or extrahealth incidents prior to the development of the neurodegenerative syndrome. Furthermore, genetic testing is not universally recommended in the clinical management guidelines of neurodegenerative diseases.8–13 As such, most neurologists, if they choose to pursue genetic testing, only screen for a small subset of genes and often choose to genotype their patients for highly penetrant and known variants rather than agnostically sequencing all neurodegenerative disease genes. Together, these common clinical ascertainment practices as well as the high costs associated with genetic testing skew the incidence rates to significantly less than what is perhaps biologically accurate. The five neurodegenerative disorders under study could partly be caused by single, rare, pathogenic variants (monogenic) or multiple, small effect variants acting synergistically to mediate disease expression (oligogenic).

1Robarts Research Institute, Schulich School of Medicine and Dentistry, Western University, London, ON, Canada; 2Department of Biochemistry, Schulich School of Medicine and Dentistry, Western University, London, ON, Canada; 3Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, ON, Canada; 4Department of Microbiology and Immunology, Faculty of Medicine, Department of Biochemistry, University of Ottawa, Ottawa, ON, Canada; 5Children’s Hospital of Eastern Ontario Research Institute, Ottawa, ON, Canada; 6Department of Clinical Neurological Sciences, Schulich School of Medicine & Dentistry, Western University, London, ON, Canada; 7Department of Medicine (Neurology), Sunnybrook Health Sciences Centre, LC Campbell Cognitive Neurology Research Unit, Hurvitz Brain Science Research Program, Sunnybrook Research Institute, University of Toronto, Toronto, ON, Canada; 8Department of Clinical Neurosciences, Cambridge Institute for Medical Research, University of Cambridge, Cambridge, UK and 9Department of Pediatrics, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada.

Correspondence: RA Hegele (hegele@robarts.ca)

Received 9 March 2016; revised 1 August 2016; accepted 5 August 2016
Advancements in next-generation sequencing (NGS) have allowed for efficient genetic variant detection at reduced costs. Currently, there are three main types of NGS applications including: (1) whole-genome sequencing (WGS); (2) whole-exome sequencing (WES); and (3) targeted gene panels.\textsuperscript{14} WGS is an indiscriminate approach that evaluates the genetic information in an individual’s entire genome. In contrast, WES targets only the protein-coding regions of the genome as disease-associated variants are significantly over-represented in coding regions.\textsuperscript{14} Consequently, WES has been one of the most widely used NGS approaches, however it still presents with several challenges. First, the cost of WES with adequate coverage (i.e., minimum × 30) still remains high at approximately $700. This makes the cumulative cost for studies with a large sample size prohibitively expensive. Second, the amount of genetic variation generated from the exome is excessive and often overwhelming for many researchers and more so for clinicians who may require the patient's genetic diagnosis to determine whether any genotype-specific treatments are available. Third, WES can generate secondary findings unrelated to the disease of interest, which should be reported to the patient’s primary healthcare provider, in accordance with the guidelines proposed by the American College of Medical Genetics.\textsuperscript{15} Thus, in both clinical and research applications, WGS or WES data are still often reduced to focus on likely pathogenic disease-specific loci. In contrast, the use of a targeted gene panel that is clinically focused on the genes underlying the disease(s) of interest, overcomes these issues that often arise when sifting through WGS and WES data.

Herein, we describe the development of a NGS based custom-designed resequencing neurodegeneration gene panel, which we have used to identify genetic variants in neurodegenerative disease cases. ‘ONDRISeq’ allows the screening of patients for variants in 80 genes implicated in neurodegenerative and cerebrovascular disease pathways. However, analysis of 80 genes can still yield an excess of genetic variation. We dichotomised all clinically relevant variants from those of uncertain significance using our integrated custom bioinformatics workflow. Our application of NGS in complex, multifactorial disorders has the potential to identify disease-specific risk markers and potentially, overlapping pathways common across all five diseases.

**RESULTS**

**Study subjects**

We recruited 216 participants affected with one of the following disorders: (1) AD/MCI, \( n = 40 \); (2) ALS, \( n = 22 \); (3) FTD, \( n = 21 \); (4) PD, \( n = 56 \); and (5) VCI, \( n = 77 \) as part of the ONDRI study (Table 1). The average age of our participants was \( 69.4 \pm 7.8 \) years. Not surprisingly, individuals diagnosed with ALS were the youngest in our cohort with an average age of \( 61.9 \pm 9.1 \) years. AD/MCI cases were the oldest patients (mean age of \( 74.5 \pm 6.6 \) years). The youngest participant in our study is a 40-year-old male diagnosed with ALS; the oldest are four 85-year-old participants (three males, one female); two diagnosed with AD/MCI and two with VCI. In general, sex ratios showed an over representation of males (male:female, 1.8:1.0), which was largely driven by the PD and VCI cases (3.3:1.0 and 2.0:1.0, respectively) similar to the known sex distribution of these disorders in prior population studies. In contrast, in the AD/MCI, ALS, and FTD cohorts, the male:female ratios did not differ considerably (1.5:1.0, 1.2:1.0; and 0.9:1.0, respectively). The self-reported ethnicity of the participants was predominantly Caucasian (82.3%) with some admixture. Overall, participants did not have a family history of neurodegenerative disease and were considered sporadic cases in our study as determined by participant recall, which was confirmed by the participant’s caregiver. Potential confounders such as age, sex, ethnicity and family history did not affect our study objectives or analysis.

**Quality assessment of ONDRISeq data**

In total, 9 independent runs of 24 samples were processed on ONDRISeq (Table 2). All targets across the 216 DNA samples were sufficiently covered (\( > \times 30 \); mean coverage \( \times 76 \pm 18 \); Table 2). On average, 22.8 million of 29.8 million reads passed quality filter equating to 77%. With the exception of the poorest performance

| Disease ID | Cases | Mean age (years ± s.d.) | Min age (years) | Max age (years) | Male: female | Self-reported ethnicity as Caucasian (%) | Family history of neurodegeneration? |
|------------|-------|-------------------------|-----------------|-----------------|--------------|------------------------------------------|-----------------------------------|
| Total      | 216   | 69.4 ± 7.8              | 40              | 85              | 140:76       | 82.3                                     | Mainly sporadic                    |
| Alzheimer's disease/mild cognitive impairment | 40 (18.5%) | 74.5 ± 6.6              | 59              | 85              | 24:16        | 93.3                                     |                                   |
| Amyotrophic lateral sclerosis | 22 (10.2%) | 61.9 ± 9.1              | 40              | 77              | 12:10        | 67.9                                     |                                   |
| Frontotemporal dementia | 21 (9.8%) | 68.8 ± 6.6              | 55              | 79              | 10:11        | 82.6                                     |                                   |
| Parkinson's disease | 56 (25.9%) | 68.0 ± 5.9              | 57              | 82              | 43:13        | 83.8                                     |                                   |
| Vascular cognitive impairment | 77 (35.6%) | 70.2 ± 7.4              | 55              | 85              | 51:26        | 84.0                                     |                                   |

**Abbreviation:** PF, passed quality filter.

**Table 2. Quality control metrics for sequencing runs on ONDRISeq**

| Parameters                  | Mean (± s.d.) | Best performance   | Poorest performance |
|-----------------------------|---------------|---------------------|---------------------|
| Cluster density (\( \times 10^7 / \text{mm}^2 \)) | 1433.6 (±165) | 1320                | 1835                |
| Target size (bp)            | 971,388       | 971,388             | 971,388             |
| Total reads (\( \times 10^6 \)) | 29.8 (±2.5)  | 29.1                | 35.6                |
| Reads PF (\( \times 10^6 \)) | 22.8 (±0.9)  | 24.1                | 22.1                |
| Reads PF (%)                | 77 (±5.8)     | 83                  | 62                  |
| Targets bases ≥ 30 (%)      | 92.0%         | 95.3                | 84.9                |
| Mean target coverage        | 76 (±18)      |                      |                     |
| Max target coverage         | 259           |                      |                     |
| Min target coverage         | 0             |                      |                     |

On average, 22.8 million of 29.8 million reads passed quality filter equating to 77%. With the exception of the poorest performance
run, all ONDRISeq runs had reads passed quality filter of >80%. Overall, 92.7% of all reads were mapped with 95% and 78% of reads mapped in the best and poorest performing runs, respectively. All other ONDRISeq runs had >90% of reads mapped. Of the matched reads, 87.1% had a Phred quality score of >30 representing a base call accuracy of 99.9%. Similarly, with the exception of the poorest performing run, all ONDRISeq runs had >85% of reads with scores >Q30. Although the poorest performing run produced lower quality data compared with the other 8 ONDRISeq runs, 84.9% of its targets were covered ≥×30 and were still analysed in our study.

Furthermore, an additional four DNA samples were extracted from brain tissue of deceased individuals. Post autopsy, sections of the brain from all four individuals were frozen for over a decade. However, we were still able to generate adequate sequence calls. Among the four samples, 96% of reads were mapped and each sample had an average coverage of ×71.

ONDRISeq is concordant with NeuroX, TaqMan allelic discrimination assay, and Sanger sequencing

We used three independent genomic techniques, NeuroX, a genome-wide array for neurodegenerative disorders, TaqMan allelic discrimination assays, and Sanger sequencing to assess the concordance with ONDRISeq in variant detection. The NeuroX array captures known polymorphic variants within the genes represented on ONDRISeq; therefore, we evaluated whether ONDRISeq could detect the same variants as NeuroX. In doing so, we processed 115 DNA samples and ONDRISeq detected all 122 non-synonymous variants initially detected by NeuroX. Furthermore, we assessed rare and common, non-synonymous and synonymous variants called by the two platforms and observed 100% concordance between calls. Of note, there were variants detected by ONDRISeq but not included on the NeuroX array. However, there were no false negatives with ONDRISeq: all variants detected by NeuroX were also detected by ONDRISeq.

Furthermore, we used a TaqMan allelic discrimination assay to genotype the same 115 DNA samples for APOE. Similarly, we observed 100% concordance between APOE genotyping calls on ONDRISeq and TaqMan.

To explore the rate of false-positive variant calls by ONDRISeq, we performed an independent concordance study for ~10% (n = 20) of randomly selected variants from samples that were called as variants by ONDRISeq using Sanger sequencing. Similar with the results of NeuroX and TaqMan allelic discrimination assay, we observed 100% concordance in variants initially detected by ONDRISeq and validated via Sanger sequencing. Thus, there were no false positives with ONDRISeq: all variants called as variants by ONDRISeq were also called as variants by validation using Sanger sequencing.

Clinical utility of ONDRISeq

All DNA samples were independently screened for a hexanucleotide expansion (G4C2) within C9orf72, a type of DNA variation that was not detectable by ONDRISeq or NeuroX. Of the 216 samples, only three (1.4%) carried an expansion within C9orf72, two were diagnosed with ALS and one with FTD (Table 3). In total, we found that only 60 out of 216 samples (27.8%) were free from rare (minor allele frequency (MAF) <1%) potentially deleterious variants (missense, nonsense, frameshift, in frame insertions and/or deletions, splicing) in ONDRISeq genes (Table 4). Of the remaining 156 cases, the AD/MCI and FTD cases had the highest variant rate based on ONDRISeq (>80%), although not necessarily disease causative. In the ALS and PD cases, we identified rare coding variants in 72.7% and 71.4% of individuals, respectively. The VCI disease cohort had the lowest number of variant carriers (65%) although still significantly higher than previous reports.16,17 Furthermore, we tabulated the number of individuals with one, two, or three or more variants. Overall, 76 (48.7%) of 156 individuals carried one variant; 57 (36.5%) carried two variants; and 23 (14.8%) carried three or more variants (Table 4).

### Table 3. Other risk variants identified in a cohort of 216 disease cases

| Disease ID | C9orf72 expansion carriers | APOE E2/E2 genotype | APOE E2/E3 genotype | APOE E2/E4 genotype | APOE E3/E3 genotype | APOE E3/E4 genotype | APOE E4/E4 genotype |
|------------|-----------------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| Total (n = 216) | 3 (1.40%) | 0 (0.00%) | 26 (12.0%) | 1 (0.46%) | 131 (60.6%) | 45 (20.8%) | 13 (6.02%) |
| AD/MCI (n = 40) | 0 (0.00%) | 0 (0.00%) | 1 (2.50%) | 0 (0.00%) | 17 (42.5%) | 15 (37.5%) | 7 (17.5%) |
| ALS (n = 22) | 2 (9.09%) | 0 (0.00%) | 4 (18.2%) | 0 (0.00%) | 12 (54.5%) | 6 (27.3%) | 0 (0.00%) |
| FTD (n = 21) | 1 (4.76%) | 0 (0.00%) | 1 (4.76%) | 0 (0.00%) | 13 (61.9%) | 5 (23.8%) | 2 (9.52%) |
| PD (n = 56) | 0 (0.00%) | 0 (0.00%) | 10 (17.9%) | 1 (1.79%) | 39 (69.6%) | 5 (8.90%) | 1 (1.79%) |
| VCI (n = 77) | 0 (0.00%) | 0 (0.00%) | 10 (13.0%) | 0 (0.00%) | 50 (64.9%) | 14 (18.2%) | 3 (3.90%) |

Abbreviations: AD/MCI, Alzheimer’s disease/mild cognitive impairment; ALS, amyotrophic lateral sclerosis; FTD, frontotemporal dementia; PD, Parkinson’s disease; VCI, vascular cognitive impairment.

### Table 4. Diagnostic yield of ONDRISeq in a cohort of 216 disease cases

| Disease ID | Individuals without any variant | Individuals with one variant | Individuals with two variants | Individuals with ≥3 variants |
|------------|--------------------------------|-----------------------------|-----------------------------|-------------------------------|
| Total (n = 216) | 60 (27.8%) | 156 (72.2%) | 76 (48.7%) | 57 (36.5%) | 23 (14.8%) |
| AD/MCI (n = 40) | 7 (17.5%) | 33 (82.5%) | 18 (54.5%) | 10 (30.3%) | 5 (15.2%) |
| ALS (n = 22) | 6 (27.3%) | 16 (72.7%) | 6 (37.5%) | 8 (50.0%) | 2 (12.5%) |
| FTD (n = 21) | 4 (19.0%) | 17 (81.0%) | 9 (52.9%) | 7 (41.2%) | 1 (5.9%) |
| PD (n = 56) | 16 (28.6%) | 40 (71.4%) | 22 (55.0%) | 13 (32.5%) | 5 (12.5%) |
| VCI (n = 77) | 27 (35.1%) | 50 (64.9%) | 21 (42.0%) | 19 (38.0%) | 10 (20%) |

Abbreviations: AD/MCI, Alzheimer’s disease/mild cognitive impairment; ALS, amyotrophic lateral sclerosis; FTD, frontotemporal dementia; PD, Parkinson’s disease; VCI, vascular cognitive impairment.

Variant criteria were based on non-synonymous, rare variants (<1% in ExAC). The variants here and in Table 5 are the same but tabulated differently.
Among the 156 cases with potentially deleterious variants, we identified a total of 266 non-synonymous, rare variants (Table 5), including 107 (40.2%) within genes known to cause the disease with which the patient has been diagnosed (e.g., variation in an AD gene in an AD patient; Table 6). An additional 159 variants (59.8%) were found in genes that were not previously associated with the respective clinical phenotype of the patient, but within a gene responsible for another disease (e.g., variation in FTD gene in an AD patient). Of the 266 variants, which will be reported on in detail upon completion of the ONDRI study of ~600 patients, 62 (23.3%) were previously reported in HGMD and/or ClinVar; whereas 204 (76.7%) were absent from disease databases (Table 5). The majority of variants not found in disease databases were observed in FTD and PD cases (88.9% and 82.6%, respectively); whereas the majority of variants present in disease databases were observed in ALS and VCI cases (35.7% and 28%, respectively; Table 5). On average, we observed four rare variants (MAF < 1%) per individual; and 1 variant per individual that met criteria set by ACMG and was considered here, as candidate variants. More rare variants were observed in individuals of African descent (16 rare variants per individual; 2 variants that met ACMG guidelines, per individual). Individuals of South Asian and Chinese origin on average carried 4.5 and 4 rare variants; and 2.5 and 2 variants meeting ACMG guidelines, respectively. These observations are likely due to ascertainment bias in the databases as they typically contain significantly more individuals of European descent than any other ethnic cohort.

Importantly, ONDRISeq is able to provide genotypes for APOE, which is not available through NeuroX and other arrays. In 216 cases, we did not identify a single case of APOE E2/E2 (Table 3). We identified 26 (12%) individuals who had an APOE E2/E3 genotype and 131 (60.6%) individuals who had an APOE E3/E3 genotype (Table 3). In total, 46 (21.3%) individuals were heterozygous for APOE E4 by possessing either an APOE E2/E4 or APOE E3/E4 genotype; whereas 13 (6.02%) individuals were homozygous for APOE E4 (Table 3). Not surprisingly, of the 13 APOE E4/E4 individuals, 7 (53.8%) were diagnosed with AD (Table 3).

Case report: strong evidence of pathogenicity for APP p.Ala713Thr in AD patient

We provide an example of a single neurodegenerative disease case to demonstrate the clinical utility of ONDRISeq and our complementary bioinformatics workflow.

The patient is a 73-year-old male diagnosed with AD. We identified a heterozygous variant, namely g.11248C>T (c.2137G>A), resulting in a missense variant p.Ala713Thr in APP, a gene known to be associated with familial autosomal dominant AD (Figure 1a). The introduction of a polar amino acid within the beta APP domain (amino-acid residues 675–713) is predicted to affect protein function according to multiple in silico analyses and generated a CADD score of 5.483 (Figure 1a,b). The affected codon is also highly conserved in evolution within the APP protein when aligned to a set of diverged species within the animal kingdom (Figure 1c). The variant is very rare with MAF of 0.006% according to Exome Aggregation Consortium (ExAC) and is absent from the 1000 Genomes database and the National Heart, Lung and Blood Institute Exome Variant Server. Furthermore, the patient is the only carrier of p.Ala713Thr in APP, among the 216 samples in our study. However, the variant has been previously observed in AD cases as it is reported in both HGMD and ClinVar databases and has been previously reported in multiple publications. Indeed, the variant had sufficient coverage of ×94, nevertheless, we independently validated the presence of the variant using NeuroX and Sanger sequencing (Figure 1a,d,e). The patient is also homozygous for APOE E3/E3.

DISCUSSION

Herein, we describe a NGS based custom-designed resequencing panel to assess genes related to neurodegenerative diseases and small vessel disease. ONDRISeq is a rapid and economical diagnostic approach that screens 80 neurodegenerative genes in parallel. We have processed a total of 216 samples on ONDRISeq in 9 runs with 24 batched samples and evaluated each run using highly stringent quality assessment criteria. With ONDRISeq, we have consistently generated high-quality data and when coupled with our bioinformatics workflow, we have been able to identify rare genetic variants in >70% of patients diagnosed with one of five diseases: AD/MCI, ALS, FTD, PD, or VCI.

The ONDRISeq calls were highly reliable based on validation by three established genetic techniques: NeuroX, a rapid and economical genome-wide genotyping-based neurodegeneration array, TaqMan allelic discrimination assay, and Sanger sequencing. Although NeuroX is able to genotype >250,000 SNPs, the advantage of ONDRISeq is that it is sequencing-based and is able to detect novel variants. This way, we can agnostically screen individuals for any novel or known variants within the 80 neurodegenerative genes. Furthermore, although the TaqMan allelic discrimination assay is a rapid genotyping approach, specific probes have to be designed for all SNPs of interest, becoming ultimately costly and inefficient. Also, unlike Sanger
| Gene                  | Chromosomal location | Affected protein                               | Associated phenotype                      | Mode of inheritance                  | OMIM numbers (focus, phenotype) |
|----------------------|----------------------|-------------------------------------------------|-------------------------------------------|--------------------------------------|---------------------------------|
| **Amyotrophic lateral sclerosis/frontotemporal dementia** |                      |                                                 |                                           |                                      |                                 |
| ALS2                 | 2q33.1               | Alsin                                           | ALS2                                      | AR (HZ), juvenile onset               | 606352, 205100                  |
| ANG                  | 1q41.1.2             | Angiogenin                                      | ALS9                                      | ADm, late onset                       | 105850, 611895                  |
| ARHgef28             | 5q13.2               | Rho guanine nucleotide exchange factor 28       | ALS and FTD                               | AR (HZ) and ADm, late onset           | 612790, PMID: 23286752 (phenotype not updated on OMIM) |
| ATXN2                | 12q24.12             | Ataxin 2                                        | ALS13                                     | ADm, late onset                       | 601517, 183090                  |
| CEP2                  | 17p11.2              | Centromere protein V                            | ALS                                       | Genetic association, late onset       | 608139, PMID: 22959728          |
| **CHMP2B**           | 3p11.2               | CHMP family member 28                           | ALS17, FTD                                | ADm, late onset                       | 609512, 614696                  |
| **DAO**              | 12q24.11             | D-amino acid oxidase                            | ALS, schizophrenia                        | ADm, late onset                       | 124050, 105400, 181500          |
| **DCTN1**            | 2p13.1               | Dynactin 1                                      | ALS, HMN78, Perry syndrome                | ADm, late onset                       | 601143, 105400, 607641, 168605  |
| Fig4                 | 6q21                 | FIG4 homologue, SAC1 lipid phosphatase domain containing | ALS11, CMT disease, YV syndrome          | ADm, late onset; AR (HZ and CH), infantile onset; AR (HZ and CH), infantile onset | 609390, 612577, 611228, 216340 |
| FUS                  | 16p11.2              | Fused in sarcoma                                | ALS6, FTD, HAT4                           | AR (HZ), ADm, late onset; AR (HZ), juvenile onset | 137070, 608030, 614782 |
| GNR                  | 17q21.31             | Granulin precursor                              | FTD, NCL                                  | ADm, late onset; ADm, early onset     | 138945, 607485, 614706          |
| **HNRNPA1**          | 12q13.13             | Heterogeneous nuclear ribonucleoprotein A1      | ALS20, inclusion body myopathy with early-onset Paget disease with/without FTD 3 | ADm, late onset; ADm, early onset | 164017, 615426, 615424 |
| **HNRNPA2B1**        | 7p15.2               | Heterogeneous nuclear ribonucleoprotein A2/B1    | Inclusion body myopathy with early-onset Paget disease with/without FTD 2 | ADm, early onset | 600124, 615422 |
| **MAPT/STH**         | 17q21.31             | Microtubule-associated protein tau              | ALS, FTD with parkinsonism, PD, AD, Pick disease, supranuclear palsy, tauopathy | ADm, late and early onset | 157140, 105400, 600274, 168600, 104300, 172700, 601104, 260540 |
| **NEFH**             | 22q12.2              | Neurofilament protein, heavy polypeptide        | ALS1                                      | ADm, late onset                       | 162230, 105400                  |
| **OPTN**             | 10p13                | Optineurin                                      | ALS12, glaucoma                           | AR (HZ) and AD, early onset           | 602432, 613435, 606657          |
| PPNL6                | 17p13.2              | Profilin 1                                      | ALS18                                     | ADm, earlier onset                    | 176610, 614808                  |
| **PRPH**             | 12q13.12             | Peripherin                                      | ALS1                                      | AR (HZ) and CH, early onset           | 603197, 612020, 215470          |
| **SETX**             | 9q34.13              | Senataxin                                       | ALS4, spinocerebellar ataxia 1            | ADm and AR, juvenile onset            | 608465, 602433, 606002          |
| SIGMAR1              | 9p13.3               | Sigma nonopiod intracellular receptor 1         | ALS16, FTD                               | ADm and AR, juvenile onset            | 601978, 614373, 105550          |
| SOD1                 | 21q22.11             | Superoxide dismutase 1                          | ALS1                                      | AR (HZ and CH), ADm, age of onset varies from 6–94 years old | 147450, 105400 |
| SQSTM1               | 5q35.3               | Sequestosome 1                                  | Paget disease of bone                     | ADm, late onset                       | 601530, 167250                  |
| **TAR15**            | 17q12                | TAF15 RNA polymerase II, TATA box-binding protein-associated factor | Paget disease of bone                     | ADm, late onset                       | 601574, 612237                  |
| **TARDBP**           | 1p36.22              | Tar DNA-binding protein                        | ALS10, FTD                                | ADm, late onset                       | 605078, 612069                  |
| **UBQLN2**           | Xp11.21              | Ubiquilin 2                                     | ALS15, FTD                                | X-linked, juvenile and late onset     | 300264, 300857                  |
| **UNC13A**           | 19p13.11             | Unc-13 homolog A (C. elegans)                   | ALS                                       | Genetic association, late onset       | 609894, PMID: 22921269 (phenotype not updated on OMIM) |
| VAPB                 | 20q13.33             | Vesicle-associated membrane protein (VAMP)-associated protein B and C | ALS, spinal muscular atrophy (Finkel type) | ADm, early and late onset             | 605704, 608627, 182980          |
| Gene   | Chromosomal location | Affected protein                                      | Associated phenotype                                                                 | Mode of inheritance        | OMIM numbers (focus, phenotype)               |
|--------|----------------------|------------------------------------------------------|--------------------------------------------------------------------------------------|---------------------------|---------------------------------------------|
| VCP    | 9p13.3               | Valosin-containing protein                           | ALS14, FTD, inclusion body myopathy with early-onset Paget disease with/without FTD | ADm, early onset          | 601023, 613954, 167320                      |
| ABCA7  | 19p13.3              | ATP-binding cassette, subfamily a, member 7          | AD                                                                                   | genetic association, late onset | 605414, 104300                              |
| APOE   | 19q13.32             | Apolipoprotein E                                     | AD2, lipoprotein glomerulopathy, sea-blue hystocye disease, macular degeneration    | AD, ADm, AR (HZ and CH), late onset | 107741, 104310, 611771, 269600, 603075      |
| APP    | 21q21.3              | Amyloid beta A4 precursor protein                    | AD 1, cerebral amyloid angiopathy                                                    | ADm and AR (HZ), early and late onset | 104760, 104300, 605714                     |
| BIN1   | 2q14.3               | Bridging integrator 1                                | AD                                                                                   | genetic association, late onset | 601248, PMID: 25365775 (phenotype not updated on OMIM) |
| CD2AP  | 6p12.3               | CD2-associated protein                               | AD                                                                                   | genetic association, late onset | 604241, PMID: 25092125 (phenotype not updated on OMIM) |
| CD33   | 19q13.41             | CD33 antigen                                         | AD                                                                                   | genetic association, late onset | 159590, PMID: 23982747 (phenotype not updated on OMIM) |
| CLU    | 8p21.1               | Clusterin                                            | AD                                                                                   | genetic association, late onset | 185430, PMID: 25189118 (phenotype not updated on OMIM) |
| CR1    | 1q32.2               | Complement component receptor 1                     | AD                                                                                   | genetic association, late onset | 120620, PMID: 25022885 (phenotype not updated on OMIM) |
| CSF1R  | 5q32                 | Colony-stimulating factor 1 receptor                 | HDLS with dementia                                                                   | ADm, early and late onset  | 164770, 221820                              |
| DNMT1  | 19p13.2              | DNA methyltransferase 1                              | HSN1E with dementia                                                                  | ADm, early onset dementia  | 126375, 614116                              |
| ITM2B  | 13q14.2              | Integral membrane protein 2B                        | Dementia                                                                            | ADm, early and late onset  | 603904, 176500, 117300                      |
| MS4A4E | 11q12.2              | Membrane-spanning 4-domains, subfamily A, member 4E | AD                                                                                   | genetic association, late onset | 608401, PMID: 21460840 (phenotype not updated on OMIM) |
| MS4A6A | 11q12.2              | Membrane-spanning 4-domains, subfamily A, member 6A | AD                                                                                   | genetic association, late onset | 606548, PMID: 21460840 (phenotype not updated on OMIM) |
| PICALM | 11q14.2              | Phosphatidylinositol-binding clathrin assembly protein | AD                                                                                   | genetic association, late onset | 650325, PMID: 24613704 (phenotype not updated on OMIM) |
| PLD3   | 19q13.2              | Phospholipase D family, member 3                     | AD19                                                                                | genetic association, late onset | 615698, 615711                              |
| PSEN1  | 14q24.2              | Presenilin 1                                         | AD3, dilated cardiomyopathy, FTD, Pick disease, acne inversa                        | ADm, early onset           | 104311, 607822, 613694, 600274, 172700, 613737 |
| PRNP   | 20p13                | Prion protein                                        | Dementia                                                                            | ADm, early onset           | 1766-40, 606688                             |
| PSEN2  | 1q32.13              | Presenilin 2                                         | AD4, dilated cardiomyopathy                                                         | ADm, early onset           | 600759, 606889, 613697                      |
| SORL1  | 11q24.1              | Sortilin-related receptor                            | AD                                                                                   | ADm, combined gene burden, late onset | 602005, 104300; PMID: 25382023 (phenotype not updated on OMIM) |
| TREM2  | 6p21.1               | Triggering receptor expressed on myeloid cells 2     | AD Nasu-Hakola disease (dementia and psychotic symptoms)                            | genetic association, late onset | 650586, PMID: 25596843 (phenotype not updated on OMIM, 221770) |
| TYROBP | 19q13.12             | Tyro protein tyrosine kinase-binding protein         | Nasu–Hakola disease (dementia and psychotic symptoms)                               | AR (HZ), juvenile onset    | 604142, 221770                             |
| Gene | Chromosomal location | Affected protein | Associated phenotype | Mode of inheritance | OMIM numbers (locus, phenotype) |
|------|----------------------|------------------|----------------------|---------------------|--------------------------------|
| Parkinson's disease | ADH1C | 4q23 | Alcohol dehydrogenase 1C, gamma polypeptide | PD, alcohol dependence protection | Genetic association, late onset | 103730, 168600, 103780 |
| | ATP13A2 (PARK9) | 1p36.13 | ATPase, type 13A2 | PD, ceroid lipofuscinosis, dementia | Genetic association, early onset and late onset | 610513, 606693 |
| | DNAK1C | 3q22.1 | DNAK/HSPA40 homolog, subfamily C, member 13 | PD | ADm, late onset | 614334, PMID: 25330418 (phenotype not updated on OMIM) |
| | EIF4G1 | 3q27.1 | Eukaryotic translation initiation factor 4-gamma | PD18 | ADm, late onset | 600495, 614251 |
| | FBXO7 | 2q12.3 | F-box only protein 7 | PD15 | AR (HZ and CH), early onset | 605648, 26300 |
| | GAK | 4p16.3 | Cyclin G-associated kinase | PD | Genetic association, late onset | 602052, PMID: 21258085 (phenotype not updated on OMIM) |
| | GCH1 | 14q22.2 | GTP cyclohydrolase I | PD, dystonia | Genetic association, early onset | 602225, 128230 |
| | GIGYF2 | 2q37.1 | GRB10-interacting GYP protein 2 | PD11 | Genetic association, early and late onset | 612003, 607688 |
| | HTRA2 | 2p13.1 | HTRA serine peptidase 2 | PD13 | ADm and genetic association, early and late onset | 606441, 610297 |
| | LRRK2 | 12q12 | Leucine-rich repeat kinase 2 | PD8 | ADm and genetic association, early and late onset | 609007, 607060 |
| | MCIR | 16q24.3 | Melanocortin 1 receptor | PD; melanoma, UV induced skin damage | Genetic association, late onset | 155555, 613099, 266300, 168600 |
| | NR4A2 | 2q24.1 | Nuclear receptor subfamily 4, group A, member 2 | PD | Genetic association, late onset | 601828, 168600 |
| | PARK2 (PRKN) | 6q26 | Pantothenate kinase 2 | Neurodegeneration | AR (HZ and CH), early onset | 601501, 600116 |
| | PARK2(DJ1) | 1q32.23 | Oncogene DJ1 | PD7 | AR (HZ and CH), juvenile onset; heterozygotes have late onset | 602533, 606324 |
| | PARL | 3q27.1 | Presenilin-associated rhomboid-like protein | PD (based on biological mechanisms, no linkage confirmed) | AR (HZ and CH), early onset | 607858, PMID: 2135049 (phenotype not updated on OMIM) |
| | PINK1 | 1p36.12 | Pten-induced putative kinase 1 | PD6 | AR (HZ and CH), ADm, early onset | 603234, 605909 |
| | PLA2G6 | 2q13.1 | Phospholipase A2, group VI | PD14, NBIA2A, NBIA2B | AR (HZ and CH), early and late onset | 603604, 612953, 256600, 610217 |
| | PM20D1 | 1q32 | Peptidase M20 domain containing 1 | PD16 | Genetic association, late onset | 603949, PMID: 25040112 (phenotype not updated on OMIM) |
| | RAB7L1 | 1q32.1 | RAB7-like 1 | PD | Genetic association, late onset | 603949, PMID: 25040112 (phenotype not updated on OMIM) |
| | SNCA | 4q22.1 | Alpha-synuclein | PD1, PD4, LBD | ADm, early onset | 163890, 168601, 60543, 127750 |
| | UCHL1 | 4p13 | Ubiquitin carboxyl-terminal esterase L1 | PD5, neurodegeneration with optic atrophy | ADm, AR (HZ), juvenile-onset | 191342, 613643, 615491 |
| | VPS35 | 16q11.2 | Vacuolar protein sorting 35 | PD17 | ADm, early and late onset | 601501, 614203 |
| | Vascular cognitive impairment | ABCG6 | 16p13.1 | ATP-binding cassette, subfamily C, member 6 | Arterial calcification; pseudoxanthoma elasticum; pseudoxanthoma elasticum forme fruste | AR (HZ), infantile onset; AR; ADm | 603234, 614473, 264800, 177850 |
| | COL4A1 | 13q34 | Collagen type IV, alpha-1 | Angiopathy, brain small vessel disease, porencephaly 1, intracerebral haemorrhage susceptibility | ADm, infantile onset | 120130, 611773, 607595, 175780, 614519 |
sequences. ONDRISeq is rapid, efficient, and economical. Following library preparation, we are able to analyse the genetic data for 24 samples in < 30 h.

We calculated the cost of sequencing 80 genes using standard Sanger sequencing. The total size of ONDRISeq is 971,388 base pairs, which can be processed via ~1,943 PCR reactions (estimation of 500 base pairs per reaction). Had we processed the sequencing reactions in bulk, the cost per sample for Sanger sequencing would have been $38,860 CND per individual. Using NGS-based approaches like WGS or WES with adequate coverage, the price still remains relatively high at $1,400 and $700 CND, respectively (prices based on The Centre for Applied Genomics, Toronto, ON, Canada; www.tcag.ca). Conversely, through strategic cost management we were able to bring our overall expenditures to a highly competitive price of $340 per sample—a reduction of >99% in cost of Sanger sequencing; a >75% reduction relative to WGS, and >50% reduction relative to WES.

Despite its efficiency and rapidity, there are still some limitations with ONDRISeq. First, it can only capture variants within the selected 80 genes, which prevents the discovery of novel disease loci. However, its custom design allows its genetic content to be altered to include novel genomic regions of interest. Second, ONDRISeq is unable to capture multi-nucleotide repeat expansions in genes, a limitation across all NGS platforms. Many neurological diseases such as Huntington’s disease, myotonic dystrophy, Friedreich’s ataxia, Fragile X syndrome, and a subset of spinocerebellar ataxias arising due to multi-nucleotide repeat expansions cannot be detected with current NGS methodologies. More recently, a hexanucleotide (G4C2) repeat expansion in C9orf72 has been observed in familial and sporadic ALS and FTD cases, and very rarely in PD cases. As such, we independently examined all repeats, which confer risk; or more than few hundred repeats and healthy population and are likely benign; 20 to few hundred alleles can range from 2 to 20 repeats, which are common in the healthy population and are likely benign; 20 to few hundred repeats, which confer risk; or more than few hundred repeats and are pathogenic. As such, we independently examined all individuals in our cohort for the C9orf72 expansion using: (1) an amplicon length PCR analysis and (2) a repeat primed PCR analysis. In doing so, we identified that 1.4% of the participants were carriers of a C9orf72 repeat expansion.

Despite these limitations and the complex heterogeneity in the five neurodegenerative diseases that are being assessed with ONDRISeq, we were able to capture rare variants with a probable, but not certain disease association based on allele frequency in the general population and the predictive score of multiple in silico software in 72.2% of cases. As the aetiology of neurodegenerative diseases is often heterogeneous and multiple factors (e.g., genetics, dietary intake, traumatic brain injury, serious infections or toxin exposure) can confer risk to disease onset, we intend to functionally validate the genetic variants, especially the novel variants, to determine their effect size and contribution to disease. Of particular interest are variants in genes with multiple disease associations as they may provide clues on the potential for development of therapy to treat symptoms common across all five neurodegenerative diseases.

**MATERIALS AND METHODS**

**Design of ONDRISeq**

Using multiple databases, we catalogued literature of neurodegeneration genetic studies. We surveyed 25 content experts (professors, scientists and clinicians within ONDRI) in molecular genetics of neurodegeneration, and used their consensus opinions to select 80 genes within the human genome that were involved in one or more of the five neurodegenerative disorders under study (Table 6). Most genes were selected based on being implicated in neurodegeneration from human genetic studies; however, some of the genes were added based on pathway analysis. Furthermore,
some genes were omitted from the ONDRISeq panel due to technical challenges, such as those involving repetitive sequence regions in the genome. This was the case for GBA gene, which is associated with an increased risk of developing PD and will thereby be assessed in separate sequencing experiments. Another gene that was omitted from the panel is C9orf72, which contains a repeat expansion and was therefore assessed with a separate genotyping assay as described in subsequent sections.

We designed a composition for detecting variants in the protein-coding regions of 80 genes summing to 1,649 targets. The 80 genes selected have a total target size of 972,388 base pairs. Using the NGS chemistry Nextera Rapid Custom Capture (Illumina, San Diego, CA, USA), we designed a total of 14,510 target speciﬁc probes that are each ~80 base pairs in length. For each library, a total of 16 pM of denatured pooled library was hybridised to target probes (two cycles of 18 h each). Samples were then amplified again to ensure speciﬁcity and greater DNA yield. A small aliquot of each library was analysed using the Agilent 2100 BioAnalyzer (Agilent Technologies, Santa Clara, CA, USA) using the ViiA 7 Real-Time PCR System (Thermo Fisher Scientiﬁc, Waltham, MA, USA) using the KAPA Quantitative PCR library quantiﬁcation kit (KAPA Biosystems, Woburn, MA, USA) and followed by subsequent serial dilutions to obtain ~5 ng/μl. Qubit 2.0 fluorometer (Thermo Fisher Scientiﬁc, Waltham, MA, USA) was then used to measure lower concentrations of DNA at a higher sensitivity.

Library preparation
Libraries were prepared in house using the Nextera Rapid Custom Capture Enrichment kit in accordance with manufacturer’s instructions. DNA samples were processed in sets of 12. DNA samples were fragmented followed by ligation of Nextera Custom Enrichment Kit-speciﬁc adapters, ampliﬁed via PCR using unique sample barcodes, equimolar pooled, and hybridised to target probes (two cycles of 18 h each). Samples were then ampliﬁed again to ensure speciﬁcity and greater DNA yield. A small aliquot of each library was analysed using the Agilent 2100 BioAnalyzer (Agilent Technologies, Palo Alto, CA, USA) to ensure adequate yield. The quantity and quality of the ﬁnal libraries were measured using the KAPA quantitative PCR library quantiﬁcation kit (KAPA Biosystems, Woburn, MA, USA) using the ViiA 7 Real-Time PCR System (Thermo Fisher Scientiﬁc).

Next generation sequencing
All samples were sequenced on the Illumina MiSeq Personal Genome Sequencer (Illumina) using the MiSeq Reagent Kit v3 in accordance with manufacturer’s instructions. Indexed samples were pooled in equimolar ratios of 500 ng. Once combined, 16 pM of denatured pooled library was
Variant calling
After demultiplexing and adapter trimming, FASTQ files were aligned to the consensus human genome sequence build GRCh37/hg19 using a customised workflow within CLC Genomics Workbench v6.5 (CLC Bio, Aarhus, Denmark) as previously described.22 Similarly, variant annotation was performed using ANNOVAR as previously described with additional databases such as CADD, HGMD (release 2015.1.), ClinVar, ExAC and our own in-house databases.

APOE genotyping
Furthermore, using ONDRISeq, in addition to screening all samples for variants within APOE, we genotyped all individuals for the APOE risk alleles rs429358 (CT) and rs7412 (CT). The combination of both individual alleles determines the APOE genotype and is known to be one of the major genetic risk factors for late onset AD.18 If there are no deletions at these loci, six potential APOE allele combinations are possible (2 alleles x 3 possible genotypes): (1) E2/E2; (2) E3/E2; (3) E4/E2; (4) E3/E3; (5) E4/E3; and (6) E4/E4, the latter of which is associated with up to an 11x increased risk in developing AD.18,36

Variant classification and prioritisation
In general, we followed the guidelines for the interpretation of sequence variants proposed by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.22 We screened for rare variants, which in our study were considered to be variants with MAF < 1% based on 1000 Genomes, NHLBI Exome Sequencing Project, and the ExAC databases. Among rare variants, we investigated whether they were non-synonymous changes (nucleotide substitutions, insertions or deletions) that resulted in missense, nonsense, splicing or frameshift variation. Variants were also assessed in silico using a compilation of prediction programs: PolyPhen-2, SIFT and CADD. HGMD and ClinVar were also integrated to determine the novelty or recurrence of any genetic variation with a specific disease state. More specifically, we were interested in determining how many variants were previously deposited into disease databases. In our study, variants were marked as clinically relevant if they were rare, resulted in non-synonymous changes, were previously observed in individuals with the same disease state, and had with similar pathogenic risks. Variants were also excluded to avoid possible genotyping error. The remaining variants were filtered to those predicted to have a potential damaging effect on protein function, according to either PolyPhen-2 or SIFT analyses implemented in ANNOVAR.

Variant validation 1: NeuroX
DNA samples were genotyped on NeuroX exome array (Illumina) according to manufacturer’s instructions. NeuroX data were loaded to GenomeStudio (Illumina) and all markers were clustered using the default Gen Call threshold (0.15); duplicate samples (N = 2) revealed identical genotypes for all markers with available genotypes (N = 268,399).22 Genotypes were converted to PLINK input files, and allele frequencies were calculated. In total, the 115 samples revealed 71,714 polymorphic autosomal markers including 43,129 exonic and 216 splicing variants; among them 39,390 polymorphisms were non-synonymous, as well as 423 stop-gain and 32 stop-loss variants, according to ANNOVAR analyses.22 Average sample call rate was 99.6%, indicating high genotype quality. Next, 1,047 polymorphic markers, which included 252 exonic variants (229 nonsynonymous and 1 splicing) within the 80 genes of the ONDRISeq targeted sequencing panel, were further processed by removing all noncoding, synonymous and common variants with MAF > 1% in any database of 1000Genomes (1000g2014oct_all), Exome Variant Server (ESP6500si_all) and ExAC. Variants overlapping segmental duplications were also excluded to avoid possible genotyping error. The remaining variants were filtered to those predicted to have a potential damaging effect on protein function, according to either PolyPhen-2 or SIFT analyses implemented in ANNOVAR.

Variant validation 2: TaqMan allelic discrimination
APOE SNP genotyping was performed using the TaqMan allelic discrimination assay for 115 samples on the 7900HT Fast Real-Time PCR System (Life Technologies, Foster City, CA, USA), and genotypes were identified using automated software (SDS 2.3; Life Technologies). Two TaqMan assays were used to determine the APOE genotype, namely (1) C_3084793_20 (rs429358: APOE codon 112) and (2) C_904973_10 (rs7412; APOE codon 158).

Variant validation 3: Sanger sequencing
Briefly, genomic DNA from the samples was first amplified via PCR, cleaned and purified, and sequenced at the London Regional Genomics Centre. Electropherograms produced were analysed using Applied Biosystems (ABI) SeqScape Software version 2.6 (Thermo Fischer Scientific, Waltham, MA, USA) with the reference sequence of each gene obtained from NCBI GenBank database.

Variant validation 4: SOD1 testing
Screening for genetic variants in the SOD1 gene was performed by PCR followed by standard Sanger sequencing methods, on DNA from four individuals diagnosed with ALS. These steps were performed in other research laboratories prior to this study. Using ONDRISeq, we sequenced DNA from these four individuals to determine whether there were any SOD1 genetic variants. This step allows us to evaluate any true/false-negative discoveries.

C9orf72 genotyping
All participants were genotyped for the G4C2-expansion in C9orf72 using a two-step method: (1) ampiclon length analysis and (2) repeat-primed PCR. Experimental procedures are described elsewhere.

Statistical analysis
The Student’s t-test was used to determine the significance of the difference among patient characteristics within the different neurodegenerative disease cohorts, where appropriate.

ACKNOWLEDGEMENTS
We thank and acknowledge the consent and cooperation of all ONDRI participants. Many thanks to the ONDRI investigators (lead investigator: MJS) and the ONDRI governing committees: executive committee; steering committee; publication committee; recruiting clinicians; assessment platforms leaders; and the ONDRI
project management team. For a full list of the ONDRI investigators, please visit: www.ONDRI.ca/people. We thank The Thunder Bay Regional Health Sciences Centre, the University of Ottawa Faculty of Medicine, and the Windsor/Essex County ALS Association. The Temerty Family Foundation provided the major infrastructure matching funds. We are grateful to Professor John Hardy from the University College London Institute of Neurology for providing valuable feedback during experimental design. SMKF is supported by the Canadian Institutes of Health Research Frederick London Institute of Neurology for providing valuable feedback during experimental matching funds. We are grateful to Professor John Hardy from the University College Association. The Temerty Family Foundation provided the major infrastructure project management team. For a full list of the ONDRI investigators, please visit: https://www.alz.co.uk/research/WorldAlzheimerReport2013.

CONTRIBUTIONS
Study design: SMKF, MG, MJ, MM, PSGH, DEB, ER and RAH; experimental procedures: SMKF, AAD, MG, EL, MZ, ADM, HC, LR and JR; data analysis: SMKF, AAD, MG, MZ and LR; manuscript preparation: SMKF, MG, PSGH, DEB, ER and RAH; study lead investigator: RAH; ONDRI lead investigator: MJ.

COMPETING INTERESTS
The authors declare no conflict of interest.

REFERENCES
1. World Alzheimer Report 2013. Journey of Caring. an Analysis of Long-Term Care for Dementia. Available at https://www.alz.co.uk/research/WorldAlzheimerReport2013.pdf. Accessed 26 August 2015.
2. WHO takes up the baton on dementia. Lancet Neurol. 14, 455 (2015).
3. Robinson, L., Tang, E. & Taylor, J. P. Dementia: timely diagnosis and early intervention. BMJ 350, h3029 (2015).
4. Hebert, L. E., Weuve, J., Scherr, P. A. & Evans, D. A. Alzheimer disease in the United States (2010-2050) estimated using the 2010 census. Neurology 80, 1778–1783 (2013).
5. Guerreiro, R., Bras, J., Hardy, J. & Singleton, A. Next generation sequencing techniques in neurological diseases: redefining clinical and molecular associations. Hum. Mol. Genet. 23, R47–R53 (2014).
6. Ronton, A. E., Chio, A. & Traynor, B. J. State of play in amyotrophic lateral sclerosis. Lancet Neurol. 14, 291–301 (2015).
7. Miller, R. G. et al. Practice parameter update: the care of the patient with amyotrophic lateral sclerosis: multidisciplinary care, symptom management, and cognitive/behavioral impairment (an evidence-based review): report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 73, 1227–1233 (2009).
8. Miller, R. G. et al. Practice parameter update: the care of the patient with amyotrophic lateral sclerosis: drug, nutritional, and respiratory therapies (an evidence-based review): report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 73, 1218–1226 (2009).
9. Grimes, D. et al. Canadian guidelines on Parkinson’s disease. Can. J. Neurol. Sci. 39, 51–30 (2012).
10. Goldman, J. S. et al. Genetic counseling and testing for Alzheimer disease: joint practice guidelines of the American College of Medical Genetics and the National Society of Genetic Counselors. Genet. Med. 13, 597–605 (2011).
11. Strong, M. J. et al. Consensus criteria for the diagnosis of frontotemporal cognitive and behavioural syndromes in amyotrophic lateral sclerosis. Amyotroph Lateral Scler. 10, 131–146 (2009).
12. Jack, C. R. Jr et al. Introduction to the recommendations from the National Institute on Aging-Alzheimer’s Association workgroups on diagnostic guidelines for Alzheimer’s disease. Alzheimers Dement. 7, 257–262 (2011).
13. Farhan, S. M. & Hegele, R. A. Exome sequencing: new insights into lipoprotein disorders. Curr. Cardiol. Rep. 16, 507 (2014).
14. Green, R. C. et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. Genet. Med. 15, 565–574 (2013).

The ONDRISeq gene panel
SMK Farhan et al.

15. Dwyer, R., Skrobot, O. A., Dwyer, J., Munafò, M. & Kehoe, P. G. Using Alzgene-like approaches to investigate susceptibility genes for vascular cognitive impairment. J. Alzheimers Dis. 34, 145–154 (2013).
16. Sun, J. H. et al. Genetics of vascular dementia: systematic review and meta-analysis. J. Alzheimers Dis. 46, 611–629 (2015).
17. Bertram, L., McQueen, M. B., Mullin, K., Blacker, D. & Tanzi, R. E. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. Nat. Genet. 39, 17–23 (2007).
18. Rossi, G. et al. A family with Alzheimer disease and strokes associated with A71T3 mutation of the APP gene. Neurology 63, 910–912 (2004).
19. Armstrong, J., Boada, M., Rey, M. J., Vidal, N. & Ferrer, I. Familial Alzheimer disease associated with A71T3 mutation in APP. Neurosci. Lett. 370, 241–243 (2004).
20. Bernardi, L. et al. AbetaPP A71T3 mutation in late onset Alzheimer’s disease with cerebrovascular lesions. J. Alzheimers Dis. 17, 383–389 (2009).
21. Ghan, M. et al. Mutation analysis of patients with neurodegenerative disorders using NeuroX array. Neurobiol. Aging 36, 545.e9–e14 (2015).
22. Singleton, A. B. Exome sequencing: a transformative technology. Lancet Neurol. 10, 942–946 (2011).
23. La Spada, A. R., Paulson, H. L. & Fischbeck, K. H. Trimetoclobide repeat expansion in neurologic disease. Ann. Neurol. 36, 814–822 (1994).
24. Stevens, J. R., Lahue, E. E., Li, G. M. & Lahue, R. S. Trimetoclobide repeat expansions catalyzed by human cell-free extracts. Cell Res. 23, 565–572 (2013).
25. DeJesus-Hernandez, M. et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 72, 245–256 (2011).
26. Renton, A. E. et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 72, 257–268 (2011).
27. Xi, Z. et al. Investigation of C9ORF72 in 4 neurodegenerative disorders. Arch. Neurol. 69, 1583–1590 (2012).
28. Xu, Z. et al. Jump from pre-mutation to pathologic expansion in C9ORF72. Am. J. Hum. Genet. 96, 962–970 (2015).
29. Sidransky, E. & Lopez, G. The link between the GBA gene andparkinsonism. Lancet Neurol. 11, 986–998 (2012).
30. Farhan, S. M. K. et al. The Ontario Neurodegenerative Disease Research Initiative (ONDRI). Can. J. Neurol. Sci (in the press).
31. Johansen, C. T. et al. LipidSeq: a next-generation clinical resequencing panel for monogenic dyslipidemias. J. Lipid Res. 55, 765–772 (2014).
32. Kircher, M. et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. 46, 310–315 (2014).
33. Landrum, M. J. et al. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. 42, D980–D985 (2014).
34. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285–291 (2016).
35. Ghebranious, N., Ivaic, L., Mallum, J. & Dokken, C. Detection of ApoE E2, E3 and E4 alleles using MALDI-TOF mass spectrometry and the homogeneous mass-extent technology. Nucleic Acids Res. 33, e149 (2005).
36. Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet. Med. 17, 405–424 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/