Investigating RFC1 Expansions in Sporadic Amyotrophic Lateral Sclerosis

Yevgenya Abramzon [1,2*], Ramita Dewan [1*], Andrea Cortese [4,5], Susan Resnick [6], Luigi Ferrucci [7], Henry Houlden [4], Bryan J. Traynor [1,2,3].

* These authors contributed equally to this work

Affiliations

1 Neuromuscular Diseases Research Section, Laboratory of Neurogenetics, National Institute on Aging, NIH, Bethesda, MD 20892, USA

2 Reta Lila Weston Institute, UCL Queen Square Institute of Neurology, University College London, London WC1N 1PJ, UK

3 Neurology Department, Johns Hopkins University, Baltimore, MD 21205, USA

4 Department of Neuromuscular Disease, UCL Queen Square Institute of Neurology, London, UK.

5 Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy

6 Laboratory of Behavioral Neuroscience, National Institute on Aging, NIH, Baltimore, MD 21224, USA

7 Translational Gerontology Branch, National Institute on Aging, NIH, Baltimore, MD 21224, USA

Corresponding author:

Yevgeniya A. Abramzon, National Institute on Aging, National Institutes of Health (NIH), 35 Convent Drive, Bethesda, MD 20892, USA. Email: yevgeniya.abramzon@nih.gov; Phone: 301-451-3550

Keywords: *RFC1*, ALS, motor neuron disorders.

Abstract

A homozygous AAGGG repeat expansion within the *RFC1* gene was recently described as a common cause of CANVAS syndrome. We examined 1,069 sporadic ALS patients for the presence of this repeat expansion. We did not discover any carriers of the homozygous AAGGG expansion in our ALS cohort, indicating that this form of *RFC1* repeat expansions is not a common cause of sporadic ALS. However, our study did identify a novel repeat conformation and further expanded on the highly polymorphic nature of the *RFC1* locus.

1. Introduction

Amyotrophic lateral sclerosis (ALS) is an invariably fatal neurological disease characterized by the progressive degeneration of upper and lower motor neurons in the brain and spinal cord. Most patients are diagnosed after 50 years of age, and worldwide prevalence will likely reach 400,000 by 2040 due to population aging (Arthur et al., 2016). Despite advances in our knowledge of the molecular and functional biology of motor neuron degeneration over the past few decades, ALS patients have few therapeutic options. The importance of understanding the genetic basis of disease in ALS is underscored by recent efforts in developing antisense oligonucleotides and other forms of precision therapies, several of which are in clinical trials and show promise of extending survival (Klim et al., 2019; McCampbell et al., 2018).

The replication factor C subunit 1 (*RFC1*) gene **(**OMIM #102579) encodes the large subunit of replication factor C, a DNA-dependent ATPase required for DNA replication and repair. The protein acts as a DNA polymerase activator by binding to the 3' end of primers and promoting the coordinated synthesis of both strands (Majka et al., 2004; Tomida et al., 2008). It may also have a role in maintaining telomere length and telomerase protein stability (Dahlen et al., 2003). Cortese and colleagues recently described a pentanucleotide AAGGG repeat expansion within intron 2 of this gene as a cause of Cerebellar Ataxia, Neuropathy, and Vestibular Areflexia Syndrome (CANVAS) (Cortese et al., 2019). We opted to examine the role of RFC1 repeat expansions in patients diagnosed with sporadic ALS, based on the phenotypic overlap between the condition and CANVAS with regard to motor neuron neuropathy, the genetic pleiotropy known to occur in the disease, and the importance of other large repeat expansions, such as C9orf72, in ALS pathogenesis.

2. Materials and Methods

2.1 Participants

We screened 1,069 American, sporadic ALS patients obtained from the Coriell repository (www.coriell.org) and 853 matched American, neurologically healthy individuals (Baltimore Longitudinal Study of Aging, USA) for the presence of repeat expansions in the *RFC1* gene (hg38, chr4:39348425-39348483). The ALS patients were diagnosed according to the El Escorial criteria (PMID: 7807156). All participants gave written informed consent, and the study complied with relevant ethical regulations. The demographic and clinical features of the cohorts are summarized in **Table 1.**

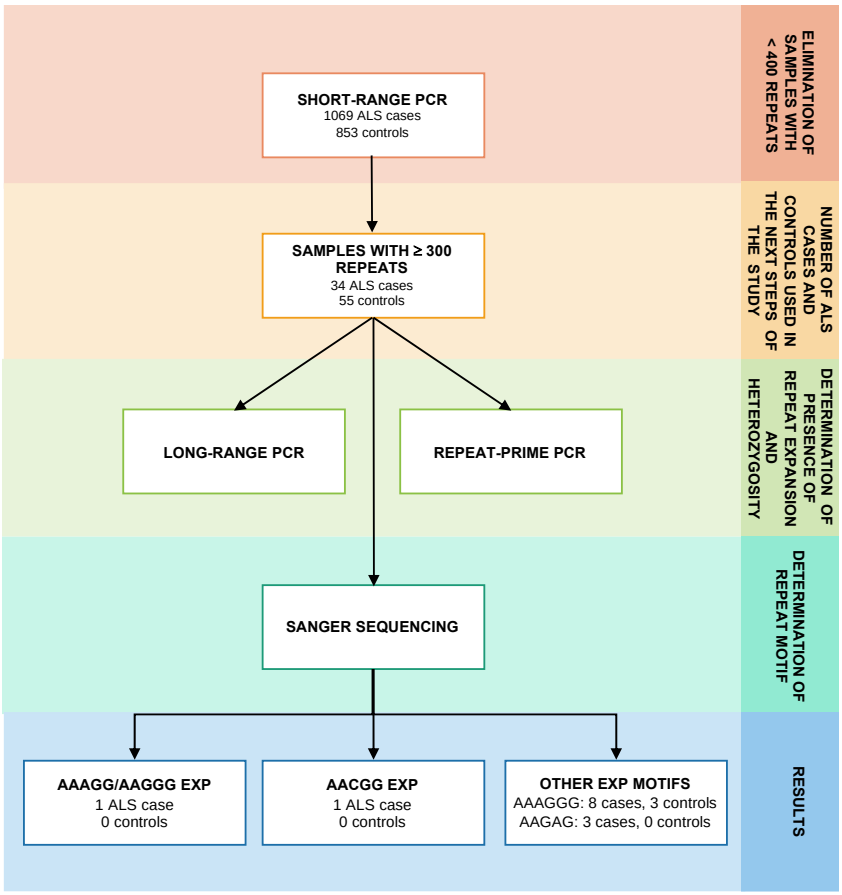
2.2 Experimental workflow

We followed the workflow described by Cortese et al to determine if a sample carried a homozygous AAGGG repeat expansion, which included a positive repeat-primed PCR for the repeat unit and a lack of PCR-amplifiable products through short-range PCR. Subsequently, amplification by long-range PCR and evaluation through Sanger sequencing was used to resolve allelic composition (Cortese et al, 2019) (**Figure 1**).

Table 1. Demographic data

Characteristic	ALS (n = 1069)	Healthy controls (n = 853)
Age at onset, years	57±12.2	N/A
Gender (male)	623 (58.3%)	467 (54.7%)
C9orf72 repeat expansion carriers	60 (5.6%)	-
Onset site		
- Bulbar	247 (23.1%)	N/A
- Limb	774 (72.4%)	N/A
- Other	48 (4.5%)	N/A

The cases were samples from the United States (n = 1069). The controls were gender and age matched samples from the United States (n = 853). All of the samples were identified as white, non-Hispanic individuals with European ancestry.



<u>Figure 1:</u> Workflow of the study.

2.3 Standard flanking PCR

DNA quality and concentration were quantified by NanoDrop™ 2000/2000c Spectrophotometer (ThermoFisher Scientific, Asheville, NC) and Qubit™ 4 Fluorometer (ThermoFisher Scientific). Standard flanking PCR primers, reagents, and cycling protocol were identical to those described by Cortese et al.(Cortese et al, 2019). PCR products were run on a 1% agarose gel (SeaKem® LE Agarose, 30 minutes at 115 volts and 500 microamperes) and analyzed for the presence of PCR-amplified products corresponding to the $[AAAAG]_{11}$ reference allele or small non-pathogenic expansions.

The standard flanking PCR protocol uses Roche FastStart™ PCR Master mix (MilliporeSigma, Burlington, MA), which can only amplify DNA segments up to 2 kilobases (kb) in length. We utilized this limitation of the standard flanking PCR to identify samples that did not amplify during this step and could potentially harbor AAGGG expansions. Samples with a band could not be carriers of homozygous large repeat expansions. A representative gel image is shown in **Figure 2**.

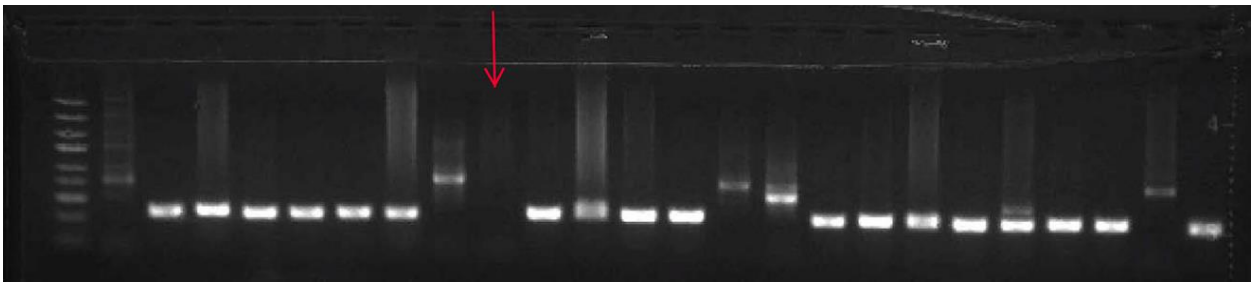


Figure 2. Standard flanking PCR demonstrating the absence of PCR-amplifiable product (marked by the red arrow). The first lane on the left contains the size reference ladder (GeneRuler 1 kb Plus DNA Ladder, Thermo Scientific). Most of the PCR fragments in this part of the study were ≤1,500 bp in length. These PCR fragments were excluded in the next steps because they accounted for less than 300 pentanucleotide repeats.

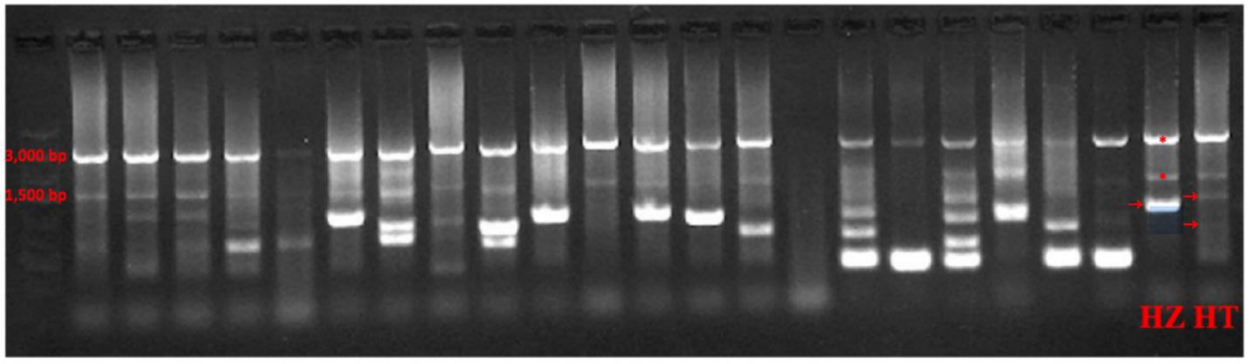
2.4 Repeat-primed PCR

Repeat-primed PCR was performed on all samples lacking PCR-amplifiable products on standard flanking PCR (n=89). The repeat-primed PCR assay was performed for each of three pentanucleotide repeat units [AAAAG/AAAGG/AAGGG] using primers, reagents, and methods, as described in Cortese et al.(Cortese et al, 2019). Repeat-primed PCR products were separated on an ABI3730xl DNA Analyzer (Applied Biosystems®, Foster City, CA). The results were visualized using GeneMapper® v.4.0 (Applied Biosystems®). Repeat-primed PCR determines if

a sample carries a repeat expansion. However, this method does not allow for accurate measurement of the repeat size of large expansions due to signal dropoff after approximately 1,000bp (equivalent to 200 repeats for a pentanucleotide motif).

2.5 Long-range PCR

Following standard flanking PCR, 89 samples in the control (n=55) and ALS (n=34) cohorts were selected for further analysis based on the absence of PCR amplifiable products (6.45% and 3.18%, respectively). PCR primers, reagents, and cycling protocol were identical to those used by Cortese et al. (Cortese et al, 2019). The long-range PCR protocol uses Phusion® High-Fidelity PCR Master Mix (New England Biolabs, Beverly, MA) and longer extension times, allowing for amplification of DNA segments up to 20 kb in length. This step allowed the determination of whether a repeat expansion present was homozygous versus heterozygous and a rough approximation of the repeat expansion size. Samples with one band were considered to harbor a homozygous repeat expansion, and samples with two distinct non-artifact bands were considered compound heterozygotes. Artifact bands were observed at ~1,500bp and 3,000bp. These bands were present in all samples and did not correspond to any previously described product size.



<u>Figure 3:</u> Long-range PCR demonstrating sample homozygosity vs. heterozygosity.

The last two lanes on the right side of the gel represented positive controls. HZ = homozygous positive control, HT = heterozygous positive control. Artifact bands (labeled with asterisk) were

observed at ~1,500bp and 3,000bp. Important bands are labeled with arrows The first lane on the left contains the size reference ladder (GeneRuler 1 kb Plus DNA Ladder, Thermo Scientific).

2.6 Sanger sequencing

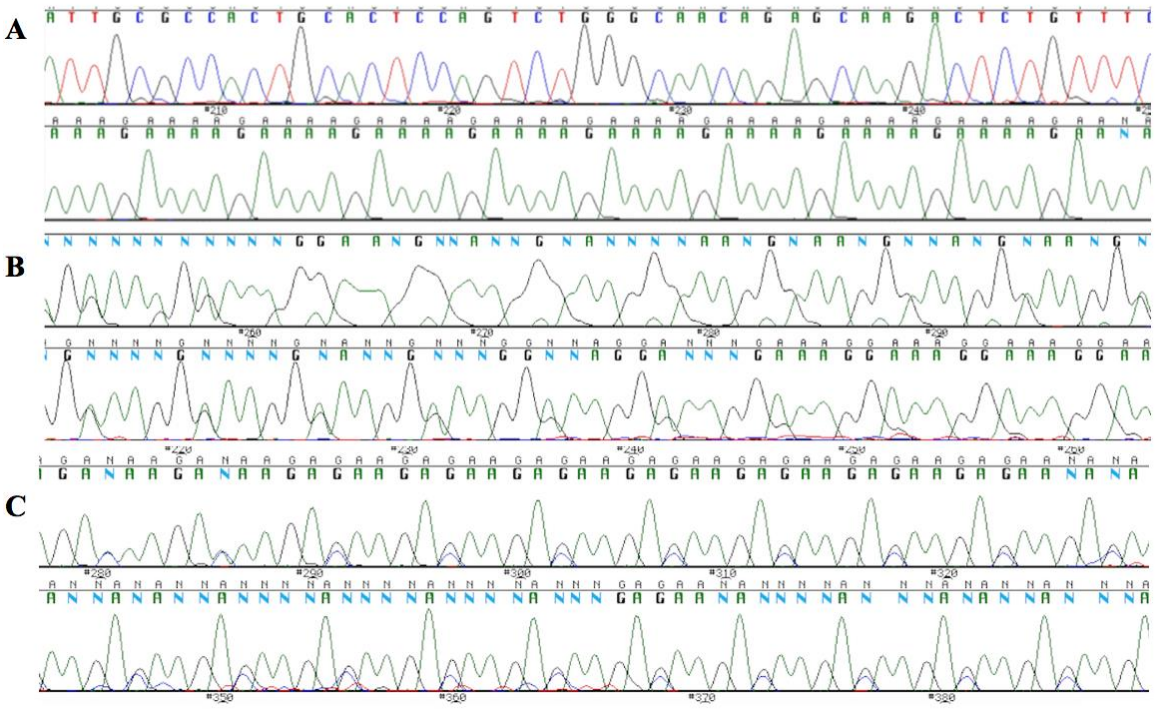
Samples that did not produce a PCR-amplifiable product during the short-range PCR step were subjected to Sanger sequencing after long-range PCR amplification to determine the allele composition. Primers were adapted from Cortese et al.(Cortese et al., 2019).

2.7 Statistical analysis

The differences between groups were assessed using the Fisher exact test.

3. Results

We did not identify the homozygous AAGGG repeat expansion, previously described as causative for CANVAS syndrome, in any of our 1,069 ALS cases or 853 control subjects. We did observe two ALS cases that were heterozygous for an enlarged $[AAGGG]_{exp}$ allele. In each case, the other allele carried an expansion in a different conformation ($[AAAGG]_{exp}$). However, this compound heterozygous genotype did not reach statistical significance for association with ALS (Fisher p-value = 0.506). Additionally, we discovered a novel *RFC1* motif, AACGG, that was present in a heterozygous state in a single ALS case. The other allele in this case was $[AAAGG]_{exp}$. The AACGG pentanucleotide motif was not found in any control individuals and was validated by Sanger sequencing (**Figure 4**). This new motif broadened the already known highly polymorphic nature of the *RFC1* locus. Additional DNA was not available from either ALS case for Southern blotting or Nanopore sequencing.



<u>Figure 4</u>. Confirmation of alternative RFC1 repeat motifs by Sanger sequencing.

A) Reference genotype $[AAAAG]_{11}/[AAAAG]_{11}$, B) heterozygous $[AAAGG]_{exp}/[AAGGG]_{exp}$ genotype, C) genotype containing the novel repeat unit $[AAAGG]_{exp}/[AACGG]_{exp}$.

We discovered several ALS cases in which the repeat expansion changed its repeat unit throughout the Sanger sequencing read (**Table 2**). These include: $[AAAGG]_n/[AAGGG]_n$ that converted to $[AAAGGG]_{exp}/[AAAGGG]_{exp}$ (controls=0.352%, ALS= 0.187%, Fisher p = 0.7); $[AAAAG]_{exp}/[AAGAG]_n$ converting to $[AAAAG]_{exp}/[AAAAG]_{exp}$ (controls=0.0%, ALS=0.094%, Fisher p = 1.0); and $[AAAAG]_{exp}/[AAAGG]_n$ converting to $[AAAAG]_{exp}/[AAAAG]_{exp}$ (controls=0.117%, ALS=0.094%, Fisher p = 1.0).

GENOTYPES	ALS		Controls	
	Number of samples	Samples %	Number of samples	Samples %
(AAGGG)exp/(AAGGG)exp	0	0	0	0
(AAAAG)exp/(AAAAG)exp	8	0.748	19	2.227
(AAAGG)exp/(AAAGG)exp	2	0.187	0	0.000
(AAAGGG)exp/(AAAGGG)exp	4	0.374	1	0.117
(AAGAG)exp/(AAGAG)exp	1	0.094	0	0.000
(AAAAG)11/(AAAAG)exp	2	0.187	4	0.469
(AAAAG)11/(AAAGGG)exp	0	0.000	1	0.117
(AAAAG)exp/(AAAGG)exp	2	0.187	1	0.117
(AAAAG)exp/(AAGAG)exp	1	0.094	0	0.000
(AAAAG)exp/(AAAGG)n(AAAAG)exp	1	0.094	1	0.117
(AAAAG)exp/(AAGGG)n(AAAAG)exp	1	0.094	1	0.117
(AAAAG)exp/(AAGAG)n(AAAAG)exp	1	0.094	0	0.000
(AAAGGG)exp/(AAGGG)n(AAAGG)exp	2	0.187	3	0.352
(AAAGG)exp/(AAGGG)exp	2	0.187	0	0.000
(AAAGG)exp/(AACGG)exp	1	0.094	0	0.000
(AAAGG)exp/(AAGGG)n(AAAGG)exp	1	0.094	0	0.000
(AAAGG)n(AAAGGG)exp/(AAGGG)n(AAAGGG)exp	2	0.187	3	0.352
(AAAGG)n(AAAGGG)exp/(AAAGG)n(AAAGGG)exp	0	0.000	2	0.234

Table 2: ***RFC1*** **genotypes in 1,069 ALS cases and 853 controls without PCR amplifiable product on flanking PCR.** Adenine bases are highlighted in red, and the cytosine bases are highlighted in blue.

Sample ID	Age of Onset	Gender	Site of Symptom	C9orf72 Status	Familial Status	EEC	Assigned Diagnosis
ND09889	54	Female	Limb-lower	WT	Sporadic	definite	ALS
ND10514	71	Male	Limb-upper	WT	Sporadic	probable	ALS
ND10311	55	Male	No Data	WT	Sporadic	probable*	PLS
ND10589	62	Male	Limb-lower	WT	Sporadic	possible*	PLS
ND04051	42	Male	Bulbar	WT	Sporadic	definite	ALS
ND12862	56	Male	Limb-lower	WT	Sporadic	definite	ALS
ND13040	70	Male	Bulbar	WT	Sporadic	probable	ALS
ND13056	48	Male	Limb-lower	WT	Sporadic	definite	ALS
ND13079	52	Female	Limb-upper	WT	Sporadic	probable	ALS
ND13135	37	Male	Limb-lower	WT	Sporadic	definite	ALS
ND13139	50	Male	Limb-upper	WT	Sporadic	probable	ALS
ND13631	50	Female	Limb-upper	WT	Sporadic	possible	ALS
ND14081	52	Male	Limb-lower	WT	Sporadic	probable	ALS
ND14122	70	Male	Limb-lower	WT	Sporadic	suspected	ALS
ND11610	46	Female	Limb-lower	WT	Sporadic	definite	ALS
ND10357	51	Female	No Data	WT	Sporadic	probable	ALS
ND07694	58	Male	Limb-lower	WT	Sporadic	definite	ALS
ND07735	41	Male	Limb-upper	WT	Sporadic	definite	ALS
ND08076	44	Male	Limb-upper	WT	Sporadic	probable	ALS
ND09115	54	Male	Bulbar	WT	Sporadic	definite	ALS
ND10556	54	Male	Bulbar	WT	Sporadic	possible	ALS
ND10557	38	Male	Limb-upper	WT	Sporadic	definite	ALS
ND11692	62	Male	Bulbar	WT	Sporadic	possibleº	PMA
ND11838	32	Male	Limb-upper	WT	Sporadic	definite	ALS
ND14729	46	Male	Limb-lower	WT	Sporadic	probable	ALS
ND14944	54	Male	Limb-lower	WT	Sporadic	definite	ALS
ND21117	73	Female	Limb-lower	WT	Sporadic	definite	ALS
ND19700	61	Male	Bulbar	WT	Sporadic	definite	ALS
ND21530	79	Female	Limb-upper	WT	Sporadic	definite	ALS
ND12167	53	Male	Limb-upper	WT	Sporadic	definite	ALS
ND12546	62	Male	Limb-lower	WT	Sporadic	probable	ALS
ND11682	63	Male	Limb-upper	WT	Sporadic	possible	ALS
ND10557	38	Male	Limb-upper	WT	Sporadic	definite	ALS
ND12546	62	Male	Limb-lower	WT	Sporadic	probable	ALS
ND12706	49	Male	Limb-lower	WT	Sporadic	suspectedº	PMA

Table 3: The phenotypes of the thirty-four patients of interest.

4. Discussion

The homozygous AAGGG repeat expansion within the second intron of the *RFC1* gene was recently reported to cause late-onset ataxia and familial CANVAS. We performed a follow-up study to examine the role of *RFC1* repeat expansion in patients diagnosed with sporadic ALS. To do this, we screened 1,069 sporadic American ALS cases and 853 US controls following the workflow described by Cortese and colleagues (Cortese et al., 2019). Homozygous *RFC1* AAGGG repeat expansions were not observed in any of our ALS cases or controls, indicating that this type of repeat expansion is not a common cause of sporadic ALS in the United States.

While our study does not support a role for the *RFC1* repeat expansion in ALS, it expands upon the dynamic nature of the *RFC1* locus. To date, four different repeat conformations have been observed in the general population: the wild type allele $[AAAAG]_{11}$, as well as longer expansions of AAAAG, AAAGG and AAGGG repeat units (Akcimen et al, 2019). Akcimen and colleagues also reported the two alternative *RFC1* repeat units AAGAG and AAAGGG (Akcimen et al, 2019). We confirmed the presence of these alternative variations in our cohorts. Furthermore, we discovered a novel *RFC1* repeat variant allele $[AACGG]_{exp}$ that was present in a single ALS case and occurred in conjunction with $[AAAGG]_{exp}$. Although the exact sizes of these alternative repeat expansions are not known due to the unavailability of additional DNA and follow-up with Southern blot, all expansions were estimated to be longer than 400 repeats based on our long-range PCR results.

Our study has limitations. First, we did not perform repeat-primed PCR, long-range PCR, and Sanger sequencing on all samples, but reserved this pipeline only for samples with no PCR-

amplifiable product on short-range PCR. Some samples may have had a single band on short-flanking PCR, and though this ruled out a homozygous repeat expansion, they may have still harbored a heterozygous expanded allele. As a consequence, we cannot determine accurate estimates of allele frequencies within our sample cohorts. Nonetheless, the novel repeat motifs are rare in our sizeable cohort consisting of over 1,000 sporadic ALS cases. Second, all of our samples were of European ancestry. It has been previously determined that the repeat expansion frequency for the same locus may vary across different populations. For example, the *C9orf72* pathogenic repeat expansion is common among European ancestry but is rare among Asian populations (Sabatelli et al, 2012; Gijselinck et al, 2012; Chio et al, 2012; Cooper-Knock et al, 2012; Majounie et al, 2012; Mok et al, 2012; Millecamps et al, 2012; Renton et al, 2011, Gromicho et al, 2018). It may be beneficial to examine regional variation in the *RFC1* locus. Third, our study focused on sporadic cases, and studies examining the role of *RFC1* expansions in familial ALS would be of interest. Fourth, our Sanger sequencing of the long-range PCR products did not cover the entire repeat length, meaning that there could have been undetected variation in that repeat expansions. Fifth, the sample size in our study was too small to rule out the possibility that homozygous (AAGGG)n RFC1 expansions are a rare cause of ALS. Sixth, the results of the short-range PCR suggest that homozygous *RFC1* locus expansions are under-represented in ALS patients (34/1069 ALS patients vs. 55/853 controls, fisher p=0.001). This is likely because most ALS cases carry less than three hundred pentanucleotide repeats on each of their two alleles.

The workflow used in this study was, by necessity, complicated due to the polymorphic nature of the *RFC1* locus. Future studies may utilize the benefits of single-molecule sequencing offered by

Pacific Bioscience (PacBio) and Oxford Nanopore Technologies (ONT). Recent studies show that long-read sequencing is well suited to characterizing known repeat expansions and discovering new disease-causing, disease-modifying, or risk-modifying repeat expansions that may have gone undetected with conventional short-read sequencing (Ebbert et al., 2018; Mitsuhashi et al., 2019). Such long-read sequencing studies in *RFC1* expansion carriers will help determine heterogeneity and its impact on ALS onset and course. Nevertheless, our study indicates that such future studies are unlikely to find that *RFC1* repeat expansions are a common cause of ALS in the United States.

Acknowledgments

The authors thank the patients and their families who donated nervous tissue for scientific research. We thank the Laboratory of Neurogenetics (NIH) staff for their collegial support and technical assistance.

Funding

This work was supported by the Intramural Research Programs of the US National Institutes of Health, National Institute on Aging (Z01-AG000949-02).

Conflict of Interest

BJT holds European, Canadian and US patents on the clinical testing for the hexanucleotide repeat expansion of *C9orf72*.

References

1. Arthur KC, Calvo A, Price TR, Geiger JT, Chio A, Traynor BJ. Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. *Nat Commun*. 2016;7:12408.
2. Talbott EO, Malek AM, Lacomis D. The epidemiology of amyotrophic lateral sclerosis. *Handb Clin Neurol.* 2016;138:225-38.
3. Renton AE, Chio A, Traynor BJ. State of play in amyotrophic lateral sclerosis genetics. *Nat Neurosci.* 2014;17:17-23.
4. Miller RG, Mitchell JD, Moore DH. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane Database Syst Rev.* 2012:CD001447.
5. Writing Group, Edaravone ALSSG. Safety and efficacy of edaravone in well defined patients with amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol.* 2017;16:505-12.
6. Klim JR, Vance C, Scotter EL. Antisense oligonucleotide therapies for Amyotrophic Lateral Sclerosis: Existing and emerging targets. *Int J Biochem Cell Biol.* 2019;110:149-53.
7. McCampbell A, Cole T, Wegener AJ, et al. Antisense oligonucleotides extend survival and reverse decrement in muscle response in ALS models. *J Clin Invest.* 2018;128:3558-67.
8. Majka J, Burgers PM. The PCNA-RFC families of DNA clamps and clamp loaders. *Prog Nucleic Acid Res Mol Biol.* 2004;78:227-60.
9. Tomida J, Masuda Y, Hiroaki H, et al. DNA damage-induced ubiquitylation of RFC2 subunit of replication factor C complex. *J Biol Chem.* 2008;283:9071-9.
10. Dahlen M, Sunnerhagen P, Wang TS. Replication proteins influence the maintenance of telomere length and telomerase protein stability. *Mol Cell Biol*. 2003;23:3031-42.

11. Cortese A, Simone R, Sullivan R, et al. Biallelic expansion of an intronic repeat in RFC1 is a common cause of late-onset ataxia. *Nat Genet.* 2019;51:649-58.
12. Akcimen F, Ross JP, Bourassa CV, et al. Investigation of the RFC1 Repeat Expansion in a Canadian and a Brazilian Ataxia Cohort: Identification of Novel Conformations. *Front Genet.* 2019;10:1219.
13. Sabatelli M, Conforti FL, Zollino M, et al. C9ORF72 hexanucleotide repeat expansions in the Italian sporadic ALS population. *Neurobiol Aging.* 2012;33:1848 e15-20.
14. Gijselinck I, Van Langenhove T, van der Zee J, et al. A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol.* 2012;11:54-65.
15. Chio A, Borghero G, Restagno G, et al. Clinical characteristics of patients with familial amyotrophic lateral sclerosis carrying the pathogenic GGGGCC hexanucleotide repeat expansion of C9ORF72. *Brain.* 2012;135:784-93.
16. Cooper-Knock J, Hewitt C, Highley JR, et al. Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. *Brain.* 2012;135:751-64.
17. Majounie E, Renton AE, Mok K, et al. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* 2012;11:323-30.
18. Mok KY, Koutsis G, Schottlaender LV, Polke J, Panas M, Houlden H. High frequency of the expanded C9ORF72 hexanucleotide repeat in familial and sporadic Greek ALS patients. *Neurobiol Aging.* 2012;33:1851 e1-5.

19. Millecamps S, Boillee S, Le Ber I, et al. Phenotype difference between ALS patients with expanded repeats in C9ORF72 and patients with mutations in other ALS-related genes. *J Med Genet.* 2012;49:258-63.
20. Renton AE, Majounie E, Waite A, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron.* 2011;72:257-68.
21. Nuytemans K, Bademci G, Kohli MM, et al. C9ORF72 intermediate repeat copies are a significant risk factor for Parkinson disease. *Ann Hum Genet.* 2013;77:351-63.
22. Lozano R, Rosero CA, Hagerman RJ. Fragile X spectrum disorders. *Intractable Rare Dis Res.* 2014;3:134-46.
23. Ebbert MTW, Farrugia SL, Sens JP, et al. Long-read sequencing across the C9orf72 'GGGGCC' repeat expansion: implications for clinical use and genetic discovery efforts in human disease. *Mol Neurodegener*. 2018;13:46.
24. Mitsuhashi S, Frith MC, Mizuguchi T, et al. Tandem-genotypes: robust detection of tandem repeat expansions from long DNA reads. *Genome Biol.* 2019;20:58.
25. Gromicho M, Pinto S, Gisca E, et al. Frequency of C9orf72 hexanucleotide repeat expansion and SOD1 mutations in Portuguese patients with amyotrophic lateral sclerosis. *Neurobiology of Aging*. 2018;70:325-e7