CLINICAL/SCIENTIFIC NOTE OPEN ACCESS

Questioning the Association of the *STMN2* Dinucleotide Repeat With Amyotrophic Lateral Sclerosis

Jay P. Ross, BSc, Fulya Akçimen, MSc, PhD, Calwing Liao, PhD, Dan Spiegelman, MSc, Ben Weisburd, BSc, Nicolas Dupré, MD, Patrick A. Dion, PhD, Guy A. Rouleau, MD, PhD, and Sali M.K. Farhan, PhD

Correspondence
Dr. Farhan
sfarhan4@gmail.com

Neurol Genet 2022;8:e678. doi:10.1212/NXG.0000000000000678

Abstract

Objectives

Recently, the number of dinucleotide CA repeats in an intron of the *STMN2* gene was reported to be associated with an increased risk for amyotrophic lateral sclerosis (ALS). Therefore, we sought to replicate this observation in an independent group of ALS patients and a much larger control group.

Methods

Here, we used whole-genome sequencing and tested the *STMN2* CA repeat in a case-control cohort of the European genetic background and in genomes from various populations in the gnomAD cohort to attempt to replicate this proposed association.

Results

We find that repeats well above the previously reported pathogenic threshold of 19 are commonly observed in unaffected individuals across different populations. Furthermore, we did not observe an association between longer *STMN2* CA repeats and ALS phenotype.

Discussion

In summary, our results do not support a role of *STMN2* CA repeats toward ALS risk. As TDP-43 aggregation is central to ALS pathogenesis, lowered expression of *STMN2* could be used as a biomarker for ALS. Therefore, a variant associated both with the risk for ALS and the level of *STMN2* expression would be clinically useful. However, for a variant to be actionable, it must be strongly replicated in independent cohorts and exceed the rigorous statistical thresholds applied.

From the Department of Human Genetics (J.P.R., F.A., C.L., G.A.R., S.M.K.F.), Montreal Neurological Institute and Hospital (J.P.R., F.A., C.L., D.S., P.A.D., G.A.R., S.M.K.F.), and Department of Neurology and Neurosurgery (D.S., P.A.D., G.A.R., S.M.K.F.), McGill University, Montréal, Quebec, Canada; Program in Medical and Population Genetics (B.W.), The Broad Institute of MIT and Harvard, Cambridge, MA; Division of Neurosciences (N.D.), CHU de Québec, Université Laval; and Department of Medicine (N.D.), Faculty of Medicine, Université Laval, Québec City, Canada.

Go to Neurology.org/NG for full disclosures. Funding information is provided at the end of the article.

The Article Processing Charge was funded by the authors.

Altered stathmin-2 (*STMN2*) expression has been implicated in amyotrophic lateral sclerosis (ALS).[1,2] On decrease of TDP-43, the *STMN2* transcript becomes truncated and produces a nonfunctional stathmin-2 protein. This dysfunction results in altered neural response to cell damage and reduced axonal regrowth.

Recently, an intronic dinucleotide CA repeat between exons 3 and 4 of *STMN2* was reported to be associated with ALS.[3] Specifically, alleles longer than 19 CA repeats were reported to increase the risk for ALS, and those carrying a 24-repeat allele alongside another long allele had the highest risk. In our study, we observed carriers of *STMN2* CA repeats well beyond the reported pathogenic repeat threshold in both case-control and gnomAD cohorts, and we did not reproduce the association between expanded *STMN2* repeats and ALS. Although *STMN2* dysfunction may contribute to ALS, its dinucleotide repeat does not impart a significant risk to ALS.

Methods

We used the STREGA checklist.[10]

Samples and Sequencing

Patients and controls were recruited in clinics across Québec, Canada. One hundred fifty-four patients (average age: 59.7 ± 11.7 years, male:female ratio 1.68) were included. Two hundred sixteen controls (average age: 67.8 ± 13.3 years, male:female ratio 0.56) were included. gnomAD was used as an external data set.[4] All individuals included gave written informed consent.

Whole-genome sequencing (WGS) was performed on Illumina HiSeq X-Ten and NovaSeq 6000 sequencers at the Génome Québec Centre d'Expertise et de Services. Bioinformatic analyses were performed on the Béluga cluster of Compute Canada and Calcul Québec using DRAGEN Bio-IT v3.8 (Illumina, Inc., San Diego, CA). After alignment to the hg38 human reference genome, an average depth of 34.1X was observed.

Estimation of *STMN2* CA Repeat Length

ExpansionHunter v4.0.2[5] was used to calculate the number of CA repeats.[3] Options applied were ReferenceRegion: chr8:79641628-79641672, VariantType: "Repeat," and LocusStructure: "(CA)*." The reported or imputed sex of individuals was used as an input option.

Statistical Analyses

Statistical tests were performed using R v4.0.3. A Fisher exact test (fisher.test) was used to test for differences between cases and controls. A Cochran-Mantel-Haenszel (CMH) test (mantelhaen.test) was used to incorporate the current results with those of the previous Theunissen study.[3] All reported *p* values are uncorrected. Per-sample allele lengths are reported in eTable 2 (links.lww.com/NXI/A735).

Data Availability

Raw data for genome sequencing used in the study are available through Project MinE (projectmine.com/) or available on request. Raw data for ExpansionHunter variant calling are also available on request.

Results

STMN2 CA repeats were successfully genotyped by ExpansionHunter in 153 ALS and 207 controls. No allele combination with the current case-control cohort suggested an association of long or long with 24 repeats (L/L with 24CA) with the ALS phenotype (Table 1). Although there was a nominally significant *p* value of L/L with 24CA using the CMH test combining allele counts from the current cohort and the Australian cohort from the previous study ($p = 0.041$), this result does not pass the multiple testing correction threshold ($\alpha = 0.05/10$; $p = 0.005$). The longest repeats were more often observed in female samples, and the largest repeats were observed in female control samples (eTable 2, links.lww.com/NXI/A735).

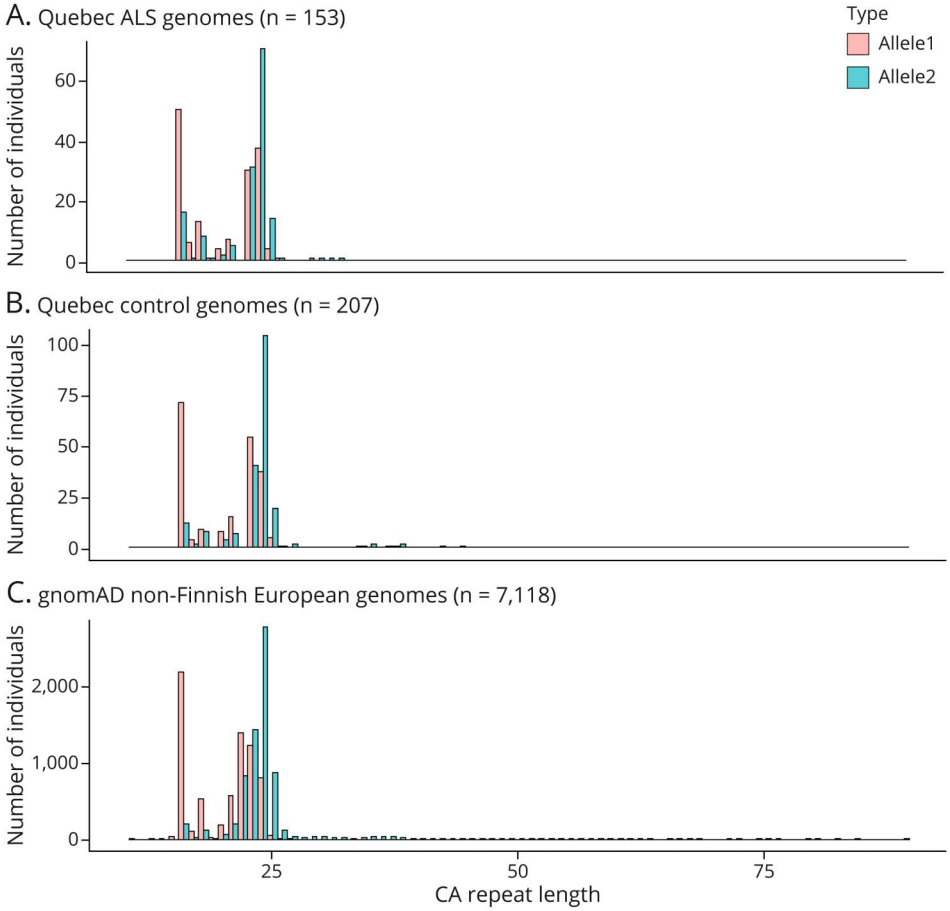
Notably, *STMN2* CA repeats much longer than the previously reported ALS-associated threshold were frequently observed in the gnomAD (eFigure 1, links.lww.com/NXI/A735).

Table 1 Replication Results of Theunissen et al.'s Associations of *STMN2* CA Repeat Lengths and ALS Phenotype

	Status		Fisher exact test			CMH test		
Genotype	**ALS**	**Controls**	***p* Value**	**OR**	**95% CI**	***p* Value**	**OR**	**95% CI**
Long/long (L/L)	84 (54.9%)	123 (59.4%)	0.4504	1.20	0.77–1.87	0.1775	1.19	0.93–1.53
L/L (with 24 CA)	59 (38.6%)	77 (37.2%)	0.8263	0.94	0.60–1.48	0.0407	1.44	1.02–2.03
L/L (without 24 CA)	25 (6.94%)	46 (12.8%)	0.1819	1.46	0.82–2.62	0.9179	0.97	0.75–1.27
Long/short (L/S)	44 (12.2%)	62 (17.2%)	0.8163	1.06	0.65–1.72	0.0935	0.79	0.61–1.03
Short/short (S/S)	25 (6.94%)	22 (6.11%)	0.1167	0.61	0.31–1.18	0.2504	1.36	0.84–2.20

Abbreviations: ALS = amyotrophic lateral sclerosis; CI = confidence interval; CMH = Cochran-Mantel-Haenszel; OR = odds ratio.
Tests of association were performed using either the Fisher exact test or the CMH test. *p* Values are reported uncorrected. Counts and percentages of individual carriers are listed for ALS and control cohorts for each combination of allele length. OR and 95% CIs are given separately for each test and combination.

Figure 1 Distributions of STMN2 CA Repeat Length Calculated From ALS (A), Control (B), and Non-Finnish European gnomAD Genomes (C)



ExpansionHunter v4.0.2 software was used to call *STMN2* CA repeat lengths from genomes aligned to the hg38 human reference genome. Sample sizes for each respective group are listed above each panel. Alleles are noted as the shorter and longer allele per individual (Allele1 and Allele2) in red and blue, respectively. ALS = amyotrophic lateral sclerosis.

Repeat lengths as long as 89 were observed in the non-Finnish European cohort, which is likely the closest match to ours and the previously reported cohort.[3] The frequency of the different allele combinations in Figure 1 varied slightly between gnomAD populations (eTable 1).

Discussion

We used WGS data to estimate the *STMN2* CA repeat length and observed large repeats above the purportedly pathogenic threshold in phenotypically normal individuals. We did not observe an association between longer alleles and ALS risk, nor did we replicate the necessity of the 24-repeat allele for this association.

The previous study reported a trend of large *STMN2* CA repeat length with decreased expression of *STMN2*.[3] However, this trend was not statistically significant. Furthermore, it is unclear whether larger repeats are linearly associated with decreased *STMN2* levels, or whether the decrease is comparable with that resulting from *TARDBP* variation or lowered *TARDBP* expression.[1,2] Although the expression level and pathologic truncation of *STMN2* are important in ALS and TDP-43 pathology, our current results refute the association of the *STMN2* CA dinucleotide repeat with ALS.

The gnomAD browser is useful to assess the maximum credible allele frequency of a variant.[7] However, as structural variants are not as well documented, it is still possible to find associations between CNVs and a given phenotype that do not replicate. Samples in gnomAD or The 1000 Genomes Project[8] may carry large repeat alleles of risk variants,[9] but without prior evidence to support variant pathogenicity, an individual might also coincidentally carry a large repeat allele. It is important that these known limitations did not hinder our evaluation of the proposed association of the *STMN2* CA repeat size and ALS.

Study Funding

J.P. Ross has received a Canadian Institutes of Health Research (CIHR) Frederick Doctoral Scholarship (FRN 159279). F. Akçimen has received funding from the Fonds de Recherche du Québec–Santé (FRQS). C. Liao has received a CIHR Vanier Graduate Scholarship (FRN 169885). P.A. Dion has received project funding from the Radala Foundation for ALS Research and jointly from the ALS Society of Canada and the Brain Canada Foundation. G.A. Rouleau has

received funding from the ALS Society of Canada and holds a Canada Research Chair in Genetics of the Nervous System and the Wilder Penfield Chair in Neurosciences.

Disclosure

The authors report no disclosures relevant to the manuscript. Go to Neurology.org/NG for full disclosures.

Publication History

Received by *Neurology: Genetics* January 26, 2022. Accepted in final form April 5, 2022. Submitted and externally peer reviewed. The handling editor was Raymond P. Roos, MD, FAAN.

Appendix Authors

Name	Location	Contribution
Jay P. Ross, BSc	Department of Human Genetics, Montreal Neurological Institute and Hospital, McGill University, Quebec, Canada	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data
Fulya Akçimen, MSc, PhD	Department of Human Genetics, Montreal Neurological Institute and Hospital, McGill University, Quebec, Canada	Analysis or interpretation of data
Calwing Liao, PhD	Department of Human Genetics, Montreal Neurological Institute and Hospital, McGill University, Quebec, Canada	Analysis or interpretation of data
Dan Spiegelman, MSc	Montreal Neurological Institute and Hospital, and Department of Neurology and Neurosurgery, McGill University, Montréal, Quebec, Canada	Analysis or interpretation of data
Ben Weisburd, BSc	Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, MA	Analysis or interpretation of data
Nicolas Dupré, MD	Division of Neurosciences, CHU de Québec, Université Laval; Department of Medicine, Faculty of Medicine, Université Laval, Québec City, Canada	Major role in the acquisition of data; additional contributions: providing clinical data
Patrick A. Dion, PhD	Montreal Neurological Institute and Hospital, and Department of Neurology and Neurosurgery, McGill University, Montréal, Quebec, Canada	Drafting/revision of the manuscript for content, including medical writing for content
Guy A. Rouleau, MD, PhD	Department of Human Genetics, Montreal Neurological Institute and Hospital, and Department of Neurology and Neurosurgery, McGill University, Montréal, Quebec, Canada	Drafting/revision of the manuscript for content, including medical writing for content
Sali M.K. Farhan, PhD	Department of Human Genetics, Montreal Neurological Institute and Hospital, and Department of Neurology and Neurosurgery, McGill University, Montréal, Quebec, Canada	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; and analysis or interpretation of data

References

1. Melamed Z, López-Erauskin J, Baughn MW, et al. Premature polyadenylation-mediated loss of stathmin-2 is a hallmark of TDP-43-dependent neurodegeneration. *Nat Neurosci*. 2019;22(2):180-190. doi: 10.1038/s41593-018-0293-z.
2. Klim JR, Williams LA, Limone F, et al. ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat Neurosci*. 2019;22(2):167-179. doi: 10.1038/s41593-018-0300-4.
3. Theunissen F, Anderton RS, Mastaglia FL, et al. Novel STMN2 variant linked to amyotrophic lateral sclerosis risk and clinical phenotype. *Front Aging Neurosci*. 2021;13:658226. doi: 10.3389/fnagi.2021.658226.
4. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. doi: 10.1038/s41586-020-2308-7.
5. Dolzhenko E, Deshpande V, Schlesinger F, et al. ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions. *Bioinformatics*. 2019;35(22):4754-4756. doi: 10.1093/bioinformatics/btz431.
6. Wickham H, Averick M, Bryan J, et al. Welcome to the Tidyverse. *J Open Source Softw*. 2019;4(43):1686. doi: 10.21105/joss.01686.
7. Whiffin N, Minikel E, Walsh R, et al. Using high-resolution variant frequencies to empower clinical genome interpretation. *Genet Med*. 2017;19(10):1151-1158. doi: 10.1038/gim.2017.26.
8. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74. doi: 10.1038/nature15393.
9. Akçimen F, Ross JP, Liao C, Spiegelman D, Dion PA, Rouleau GA. Expanded CAG repeats in ATXN1, ATXN2, ATXN3, and HTT in the 1000 Genomes Project. *Mov Disord*. 2021;36(2):514-518. doi: 10.1002/mds.28341.
10. Little J, Higgins JP, Ioannidis JP, et al; STrengthening the REporting of Genetic Association Studies. STrengthening the REporting of Genetic Association Studies (STREGA): an extension of the STROBE statement. *PLoS Med*. 2009;6(2):e22. doi: 10.1371/journal.pmed.1000022.