EDITORIAL OPEN ACCESS

Intronic pentanucleotide expansion in the replication factor 1 gene (*RFC1*) is a major cause of adult-onset ataxia

Sylvia M. Boesch, MD, and Martha A. Nance, MD

Neurol Genet 2020;6:e436. doi:10.1212/NXG.0000000000000436

Correspondence
Ass. Prof. Dr. Boesch
sylvia.boesch@i-med.ac.at

RELATED ARTICLE

Prevalence of RFC1-mediated spinocerebellar ataxia in a North American ataxia cohort

Page e440

The ataxias comprise diseases of both genetic and nongenetic origin with extreme clinical and genetic heterogeneity. They may present as a pure cerebellar form or as part of a more complex neurologic syndrome. Progressive, neurodegenerative sporadic adult-onset ataxias (SAOAs) without a known cause have a prevalence rate of 2.2–12.4 per 100,000. In several ataxia cohorts, repetitive genetic screening using high-coverage ataxia-specific gene panels in combination with next-generation sequencing (NGS) failed to identify a causative gene in 50%–90% of SAOAs.[1–3] Cerebellar ataxia, neuropathy, vestibular areflexia syndrome (CANVAS), first described by Brownstein et al.,[4] is a slowly progressive neurodegenerative disorder with adult onset, affecting the cerebellum, sensory neurons, and the vestibular system. CANVAS is usually sporadic, but occasionally occurs in siblings. Two research groups recently identified large biallelic intronic AAGGG expansions in replication factor C subunit 1 (*RFC1*) resulting in CANVAS, an adult-onset neurodegenerative ataxia.[5–7] *RFC1* normally loads proliferating cell nuclear antigen onto DNA and activates DNA polymerases δ and ε to promote the coordinated synthesis of both strands during replication or after DNA damage.[8]

In this issue of *Neurology® Genetics*, Syriani et al.[9] investigated the prevalence of intronic AAGGG expansions in the *RFC1* gene in a North American cohort of 911 predominantly adult-onset patients with undiagnosed familial or sporadic ataxia. Testing in this cohort revealed 29 patients with biallelic expansions (3.2%), one-third of whom had the full CANVAS syndrome. The remaining had late-onset ataxia frequently accompanied by neuropathy (60%). All *RFC1* expansion carriers were Caucasian. The rate of heterozygosity was as high as 6.8%, which may be caused by overrepresented alleles with repeat lengths below 400 repeats—the pathogenic threshold in *RFC1* anticipated in 2 previous studies.

Nucleotide repeat disorder reloaded

Repetitive DNA sequences constitute approximately one-third of the genome. There is evidence that they may contribute to diversity within and between species. They display considerable variability in length between individuals, which is presumed to have no detrimental consequences unless the repeat number is expanded beyond a gene-specific threshold. Pathologic unstable repeat expansions are classified according to their length, repeat sequence, gene location, and underlying pathologic mechanisms. Large (hundreds-thousands of copies) pathogenic repeat expansions are typically located in noncoding regions including promoters, introns, and untranslated regions of genes and can show somatic instability. Repeat expansions in introns are thought to produce aberrant repeat-bearing RNAs that interact with and sequester a wide variety of essential proteins, resulting in cellular toxicity.

Targeted non–sequence-based testing is still the method of choice to detect nucleotide repeat expansions in the human genome. Commonly used NGS techniques such as whole genome sequencing (WGS) and whole exome sequencing (WES) fail to detect repetitive regions.

From the Medical University Innsbruck (S.M.B.); and Struthers Parkinson's Center (M.A.N.).

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

Indeed, in the initial study in inherited CANVAS, WGS did not detect the causative mutation in *RFC1*.[5] The use of targeted non–sequence-based techniques and Southern blot finally led to the detection of 4 distinct intronic repeat conformations in *RFC1*: $AAAAG_{11}$ (the wild-type sequence) and longer expansions of $AAAAG_n$, $AAAGG_n$, and $AAGGG_n$. Of these, the AAGGG pentanucleotide expanded up to 400–2,000 repeats was the only disease-causing condition. The expansion occurs in the poly(A) tail of an AluSx3 element, and differs in both size and nucleotide sequence from the reference $(AAAAG)_{11}$ allele.

Haplotype and allele carrier frequency in *RCF1*

The same ancestral haplotype is shared by the majority of familial and positive *RFC1* cases, as well as some healthy carriers of 2 $(AAAGG)_{exp}$ alleles. It is likely that the nucleotide change from AAAAG to AAAGG or AAGGG represents an ancestral founder event, followed by the pathologic expansion of the repeated unit, whose size seems to be related to its guanine-cytosine content. Up to now, analyses of the core haplotype in the mixed ethnic cohort confirmed the European core haplotype estimated to have arisen more than 25,000 years ago. Although the AAGGG repeat expansion has been identified in non-European individuals (Native American, Arabic, and Japanese), it remains highly overrepresented in populations of European descent, with frequencies of 4%–6.8% (White and Hispanic).[7,9–11]

In summary, the study by Syriani et al.[9] supports the notion that the newly discovered *RCF1* gene is a major cause of CANVAS. Analysis of *RCF1* should be included in clinical diagnostic testing of adult-onset neurodegenerative ataxia, especially when neuropathy is present. There are multiple areas for future work, including deep phenotyping in sporadic adult-onset ataxias, analyses for a correct determination of pathogenic repeat lengths, and the stability of this pentanucleotide repeat sequence across siblings and generations within families. Moreover, an understanding of the fine structure of *RFC1* as it relates to the final repeat composition of the pathogenic pentanucleotide and its function remains to be elucidated by additional studies.[5,6,10]

The fact that this highly prevalent ataxia gene was unknown until now, highlights both the importance of precise phenotyping and sampling, as well as the use of analytic techniques beyond currently available panels and NGS. Intronic repeat expansions, in particular, are difficult to identify but may be common causes of neurodegenerative disease.

Study funding

No targeted funding reported.

Disclosure

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

Publication history

Received by *Neurology: Genetics* April 15, 2020. Accepted in final form April 20, 2020.

References

1. Fogel BL, Lee H, Deignan JL, et al. Exome sequencing in the clinical diagnosis of sporadic or familial cerebellar ataxia. JAMA Neurol 2014;71:1237–1246. Erratum in: JAMA Neurol 2015;72:128.
2. Giordano I, Harmuth F, Jacobi H, et al. Clinical and genetic characteristics of sporadic adult-onset degenerative ataxia. Neurology 2017;89:1043–1049.
3. Sun M, Johnson AK, Nelakuditi V, et al. Targeted exome analysis identifies the genetic basis of disease in over 50% of patients with a wide range of ataxia-related phenotypes. Genet Med 2018;21:195–206.
4. Bronstein AM, Mossman S, Luxon LM. The neck-eye reflex in patients with reduced vestibular and optokinetic function. Brain 1991;114:1–11.
5. Cortese A, Simone R, Sullivan R, et al. Author Correction: biallelic expansion of an intronic repeat in RFC1 is a common cause of late-onset ataxia. Nat Genet 2019;51:920.
6. Cortese A, Tozza S, Yau WY, et al. Cerebellar ataxia, neuropathy, vestibular areflexia syndrome due to RFC1 repeat expansion. Brain 2020;143:480–490.
7. Rafehi H, Szmulewicz DJ, Bennett MF, et al. Bioinformatics-based identification of expanded repeats: a non-reference intronic pentamer expansion in RFC1 causes CANVAS. Am J Hum Genet 2019;105:151–165.
8. Overmeer RM, Gourdin AM, Giglia-Mari A, et al. Replication factor C recruits DNA polymerase delta to sites of nucleotide excision repair but is not required for PCNA recruitment. Mol Cell Biol 2010;30:4828–4839.
9. Syriani DA, Wong D, De Gusmao CM, et al. Prevalence of RFC1-mediated spinocerebellar ataxia in a North American ataxia cohort. Neurol Genet 2020;6:e440. doi: 10.1212/NXG.0000000000000440.
10. Akçimen 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.
11. Nakamura H, Doi H, Mitsuhashi S, et al. Long-read sequencing identifies the pathogenic nucleotide repeat expansion in RFC1 in a Japanese case of CANVAS. J Hum Genet 2020;65:475–480.