RFC1-Related Disease
Molecular and Clinical Insights

Kayli Davies, BSc Hons, David J. Szmulewicz, MBBS (Hons), PhD, FRACP, Louise A. Corben, PhD, Martin Delatycki, MD, PhD, and Paul J. Lockhart, PhD

Neurol Genet 2022;8:e200016. doi:10.1212/NXG.000000000000200016

Abstract

In 2019, a biallelic pentanucleotide repeat expansion in the gene encoding replication factor C subunit 1 (RFC1) was reported as a cause of cerebellar ataxia with neuropathy and vestibular areflexia syndrome (CANVAS). In addition, biallelic expansions were shown to account for up to 22% of cases with late-onset ataxia. Since this discovery, the phenotypic spectrum reported to be associated with RFC1 expansions has extended beyond the initial conditions to include pure cerebellar ataxia, isolated somatosensory impairment, combinations of the 2, and parkinsonism, leading to a potentially broad differential diagnosis. Genetic studies suggest RFC1 expansions may be the most common genetic cause of ataxia and are likely underdiagnosed. This review summarizes the current molecular and clinical knowledge of RFC1-related disease, with a focus on the evaluation of recent phenotype associations and highlighting the current challenges in clinical pathways to diagnosis and molecular testing.
The replication factor C subunit 1 (RFC1) gene encodes the large subunit of replication factor C, a pentameric DNA polymerase accessory protein involved in DNA replication and repair.1 Recently, a biallelic AAGGG repeat expansion in intron 2 of RFC1 was identified as a cause of cerebellar ataxia with neuropathy and vestibular areflexia syndrome (CANVAS) and other phenotypes of late-onset cerebellar ataxia (CA).2,3 This pathogenic expansion differs in both repeat motif and size from the reference AAAAG(11) allele. Since its identification, estimates of the carrier frequency of the pathogenic AAGGG expansion have ranged from 0.7% to 6.5% in different control populations (Table 1).2-4 This variability likely reflects the modest cohort sizes and ethnicities examined to date; we anticipate accurate population-specific carrier frequency estimates will be available when bioinformatic tools that can identify the presence of the pathogenic repeat expansion in NGS data5 are applied to large population databases such as Genome Aggregation Database. Multiple studies have identified biallelic RFC1 expansions in clinical cohorts with ataxia, ranging from 1% to almost 30% of affected individuals (Table 2). Stratification of these cohorts suggests a high yield in cases meeting the clinical criteria of CANVAS, a variable yield for phenotypes that have 1 or more of CA and somatosensory impairment or bilateral vestibulopathy and a low yield for individuals with isolated CA. While generalizing the contribution of expansions in RFC1 to disease remains challenging due to variability in the inclusion criteria for cohort studies, the results to date clearly demonstrate that RFC1 expansions are potentially the most common genetic cause of CA and are likely underdiagnosed. Current molecular diagnosis of RFC1 expansion requires PCR-based testing for flanking sequence and repeat-primed PCR using motif-specific primers. These assays are generally performed in a research setting, with limited diagnostic testing currently available. Notably, the phenotypic spectrum associated with RFC1 expansions has broadened beyond the classical CANVAS presentation. Allowing for the fact that dysautonomia and a chronic cough was described in the early CANVAS literature,6,7 biallelic RFC1 expansions have been reportedly associated with ataxia of purely cerebellar origin,8,9 isolated somatosensory (because both peripheral neuropathy and neuronopathy [ganglionopathy] have been reported in RFC1-related disease, the term somatosensory, which covers both pathologies, is used) impairment,10-12 and combinations of the 2,9,13-16 in addition to parkinsonism.17-21 Therefore, it can be difficult to both decide who is appropriate to test and where to source molecular testing. The size of the expanded repeat can be estimated by Southern blot analysis or long-read sequencing,22 but neither method is currently offered in a diagnostic setting. Moreover, detailed knowledge of the pathogenic size threshold is still lacking, with current pathogenic estimates ranging from 400 to several thousand repeats. In addition to addressing the deficits in molecular diagnostic tools, progress in RFC1-related disease clinical management and diagnostics requires advancing beyond the bedside examination of potentially affected individuals and exploiting objective metrics. The purpose of this review was to highlight the current genotypic and phenotypic knowledge of RFC1-related disease. A literature search of PubMed from the year CANVAS was first defined to current (January 2011 to December 2021) was performed using the terms ataxia OR CANVAS OR RFC1, with the results manually curated for relevance. If appropriate, primary sources quoted in references were also retrieved and curated. Pathologic descriptions were limited to studies that performed molecular confirmation of biallelic RFC1 expansion. Our aim was to raise awareness of the spectrum of disease that can be caused by RFC1 repeat expansions and highlight the importance of deep phenotyping for furthering our understanding of the spectrum of RFC1-related disease.

Gene Discovery and Molecular Genetics

CANVAS was first defined as a clinical entity in 2011 after deep and careful phenotyping identified a somatosensory abnormality as a feature of many individuals with the combination of CA with bilateral vestibulopathy.5,23 Greater than a third of affected individuals with a clinical diagnosis of CANVAS identified in these early studies had an affected family member, most of whom were siblings, suggesting a potential recessive mode of inheritance.6,24 Despite the presumed genetic basis for the condition, the challenges expanded repeats present for traditional molecular biology tools such as PCR and cloning meant the repeat expansion in RFC1 was only recently identified as the cause of CANVAS. This discovery was made by 2 groups after the identification of a linkage region on chromosome 4p14 in multiple families with a clinical diagnosis of CANVAS.2,5 Using different approaches to analyze whole-genome sequencing (WGS) data, both groups subsequently identified the nonreference AAGGG(n) expansion in RFC1 within the linkage region. One team identified the expansion by visually inspecting the aligned WGS read pairs within the linkage region, which revealed a region of reduced read depth in all patients with CANVAS.
corresponding to the repeat in RFC1. By contrast, the second team used a bioinformatics-based approach to identify repeat expansions genome-wide resulting in the identification of the AAGGG(n) repeat within the linkage interval. Clinically, CANVAS is the triad of (1) cerebellar impairment, (2) bilateral vestibular hypofunction, and (3) somatosensory abnormalities.\(^2,3\) The clinical diagnosis of CANVAS is made using detailed oculomotor, electrophysiologic, and imaging criteria (Table 3), which were developed after careful phenomenotyping of affected individuals.\(^2,5\) This deep phenomenotyping and development of diagnostic criteria proved an important step in the eventual identification of the causative gene. After the

### Table 1 Carrier Frequency of Pathogenic AAGGG(n) Allele in Previously Reported Cohorts

<table>
<thead>
<tr>
<th>Studies</th>
<th>Ethnicity</th>
<th>Cohort</th>
<th>AAGGG(_{\text{\text{n}}}) frequency</th>
<th>Method of identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortese et al., 2019(^2)</td>
<td>European</td>
<td>304 controls</td>
<td>0.7%</td>
<td>PCR</td>
</tr>
<tr>
<td>Rafehi et al., 2019(^3)</td>
<td>European</td>
<td>31 controls</td>
<td>6.5%</td>
<td>Bioinformatics</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>69 Coriell WGS</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>133 GTEx WGS</td>
<td></td>
</tr>
<tr>
<td>Akcimen et al., 2019(^6)</td>
<td>Canadian</td>
<td>163 controls</td>
<td>4.0%</td>
<td>PCR</td>
</tr>
<tr>
<td>Fan et al., 2020(^6)</td>
<td>Chinese Han</td>
<td>245 controls</td>
<td>2.2%</td>
<td>PCR</td>
</tr>
<tr>
<td>Wan et al., 2020(^9)</td>
<td>Chinese</td>
<td>203 controls</td>
<td>1.0%</td>
<td>PCR</td>
</tr>
</tbody>
</table>

Abbreviations: GTEx = Genotype-Tissue Expression project; n.d. = no data; WGS = whole-genome sequencing.

### Table 2 Frequency of Biallelic RFC1 Expansions

<table>
<thead>
<tr>
<th>Studies</th>
<th>Total cohort (%)</th>
<th>CANVAS (%)</th>
<th>CA + BV or PN (%)</th>
<th>CA(^b) (%)</th>
<th>MSA (%)</th>
<th>PN (%)</th>
<th>CMT (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cortese et al., 2019(^2)</td>
<td>33/150 (22.0)</td>
<td>11/12 (91.1)</td>
<td>21/39 (53.8)</td>
<td>1/99 (1.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akcimen et al., 2019(^6)</td>
<td>2/177 (1.1)</td>
<td>2/177 (1.1)</td>
<td>1/177 (0.1)</td>
<td>0/111 (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cortese et al., 2020(^3)</td>
<td>205/363 (28.9)</td>
<td>63/70 (90.0)</td>
<td>42/293 (14.3)</td>
<td>0/336 (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sullivan et al., 2020(^1)</td>
<td>29/911 (3.2)</td>
<td>8/11 (72.7)</td>
<td>14/63 (22.2)</td>
<td>7/726 (1.0)</td>
<td>0/111 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tsuchiyu et al., 2020(^4)</td>
<td>4/37 (10.8)</td>
<td>4/4 (100)</td>
<td>0/33 (0)</td>
<td>4/37 (10.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Van Dele et al., 2020(^5)</td>
<td>9/85 (10.6)</td>
<td>3/33 (100)</td>
<td>6/29 (20.7)</td>
<td>5/33 (15.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gisatulin et al., 2020(^6)</td>
<td>21/96 (21.9)</td>
<td>15/17 (88.2)</td>
<td>2/9 (22.2)</td>
<td>4/70 (5.7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wan et al., 2020(^9)</td>
<td>4/386 (1.0)</td>
<td>1/104 (1.0)</td>
<td>3/282 (1.1)</td>
<td>0/336 (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traschutz et al., 2021(^3)</td>
<td>70/181 (38.7)</td>
<td>66/88 (75.0)</td>
<td>70/181 (38.7)</td>
<td>0/336 (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sullivan et al., 2021(^6)</td>
<td>15/163 (9.2)</td>
<td>14/18 (77.8)</td>
<td>1/103 (1.0)</td>
<td>0/42 (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Montaut et al., 2021(^8)</td>
<td>5/77 (6.5)</td>
<td>4/5 (80.0)</td>
<td>0/10 (0)</td>
<td>0/162 (0.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kontogeorgiou et al., 2021(^9)</td>
<td>34/234 (14.5)</td>
<td>34/234 (14.5)</td>
<td>34/234 (14.5)</td>
<td>34/234 (14.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curro et al., 2021(^1)</td>
<td>43/225 (19.1)</td>
<td>12/12 (100)</td>
<td>13/13 (100)</td>
<td>18/200 (9.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lipponen et al., 2021(^1)</td>
<td>9/150 (6.0)</td>
<td>5/5 (100)</td>
<td>0/91 (0)</td>
<td>4/54 (7.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: BV = bilateral vestibulopathy; CA = cerebellar ataxia; CANVAS = cerebellar ataxia, neuropathy, and vestibular areflexia syndrome; CMT = Charcot-Marie-Tooth; MSA = multiple system atrophy; PN = peripheral neuropathy.

\(^a\) The authors have attempted to unify the cohorts into common ataxia subgroups (CANVAS, cerebellar ataxia and bilateral vestibulopathy or peripheral neuropathy and cerebellar ataxia) based on the information provided in each study. The diagnoses of MSA, CIAP, and CMT were made by authors of the respective studies. As described in this review, in some cases, we do not believe that sufficient evidence has been provided to support some of these classifications.

\(^b\) CA with no other features of CANVAS or status of other features unclear.

\(^c\) Pathologically confirmed MSA.

\(^d\) Only 3 of the 9 positive cases had vestibular testing, all 3 showed vestibular hypofunction and are therefore classified as CANVAS.
identification of the pathogenic AAGGG\(_n\) allele, genetic analysis identified a core ancestral haplotype that most likely arose in Europe 25,000 years ago, suggesting a single founder event for the pathogenic repeat expansion.\(^2,3\)

The RFC1 repeat locus is highly polymorphic in nature with considerable variability in both repeat motif and size in the population (Figure 1). In the first study to examine the RFC1 locus in unaffected control populations, 4 different repeat conformations were observed: the reference AAAAG\(_n\), nonpathogenic expansions of AAAAG\(_n\) and AAAGG\(_n\) motifs, and the pathogenic expansion of AAGGG\(_n\).\(^2\) Nonpathogenic expanded alleles account for approximately 20% of alleles in unaffected individuals.\(^2\) These nonpathogenic expanded alleles range in size from 40 to 1,000 repeats for the AAAGG\(_n\) motif and from 15 to 200 repeats for the AAAAG\(_n\) motif. Initial studies suggested expansions of AAGGG\(_n\) exceeding \(\sim 400\) repeats are pathogenic, while the non-pathogenic expansions of both AAAAG\(_n\) and AAAGG\(_n\) motifs are typically smaller in size.\(^2\) Three additional motifs have subsequently been described in the general population: AAGAG\(_n\), AGAGG\(_n\),\(^26\) and most recently, AACGG\(_n\).\(^27\) However, no individuals homozygous for these conformations have been reported, so the potential pathogenicity of these motifs remains uncertain. In addition to the AAGGG\(_n\) allele, 2 alternate pathogenic allele configurations have been identified in Asia Pacific cohorts. An apparently Māori-specific allele configuration of AAAGG\(_{10-25}\)AAGGG\(_n\)AAAGG\(_{4-6}\) was identified in individuals affected by CANVAS with New Zealand and Cook Island Māori heritage; affected individuals were phenotypically indistinguishable from AAGGG\(_n\)-mediated disease.\(^28\) In addition, a novel pathogenic ACAGG\(_n\) motif was also identified in both Asia Pacific and Japanese cohorts.\(^29,30\) Additional phenotypic features beyond the classical CANVAS

**Figure 1** Polymorphism at the RFC1 Locus

![RFC1 Polymorphism Diagram](image-url)

Schematic representation of the repeat configurations at the RFC1 repeat locus showing the nonpathogenic alleles (top), pathogenic alleles (middle), and alleles of uncertain pathogenicity (bottom). RFC1 = replication factor C subunit 1.
triplet have been reported in affected individuals with expansions of this ACAGG(n) motif, including fasciculations and elevated serum creatine kinase levels.\textsuperscript{29} Notably, alleles with both the ACAGG(n) and AAAGG(10-25)AAGGG(n)AAAGG(4-6) configurations share the same core haplotype as the AAGGG(n) motif.\textsuperscript{28,29} While the results of the 4 haplotype studies performed to date are suggestive,\textsuperscript{2,3,28,29} additional analyses across additional populations are required to confirm the theory of a single founder event for the pathogenic expansion.\textsuperscript{3}

Reported sizes for the pathogenic AAGGG(n) alleles are typically larger, ranging from \(\sim 400\) to \(2,000\) repeats. However, expansions of \(\sim 100-160\) repeat units have recently been reported as likely pathogenic in 4 individuals with a clinically defined movement disorder (see further).\textsuperscript{19} Of interest, and in contrast to other repeat expansion conditions such as Friedreich ataxia and CAG expansion disorders, no significant correlation between the size of the repeat expansion and disease severity or age at onset has been reported to date.\textsuperscript{13} This lack of correlation together with the high frequency of nonpathogenic expanded alleles supports the hypothesis that the repeat motif rather than repeat length may be the driver of pathogenicity. This hypothesis is supported by findings in other pentanucleotide repeat expansion disorders, including spinocerebellar ataxia (SCA) 31, SCA37, and familial adult myoclonic epilepsy, where only expansions of a nonreference motif.\textsuperscript{28,29} In addition, interruptions of an ATCCT motif in SCA10 are strongly associated with epileptic seizures, providing additional evidence that the repeat motif rather than length is driving pathogenicity.\textsuperscript{34} Similar to RFC1, nonpathogenic alleles at the SCA37 locus are highly polymorphic, suggesting there may be a common mechanism driving pathogenicity in pentanucleotide repeat disorders.\textsuperscript{32} Further studies are therefore needed to accurately elucidate the pathogenic expansion threshold and the contribution of various motifs to pathogenesis.

**Neuropathology**

Although the pathomechanisms of RFC1 expansions remain unknown, it is well established that neuronal loss occurs in ataxia. Current knowledge of the neuropathology of RFC1 expansion is limited and derives from postmortem examinations of 5 individuals with a clinical diagnosis of CANVAS and confirmed biallelic AAGGG expansions.\textsuperscript{2,9,23,35,36} Postmortem histopathologic examination of brain tissue from these individuals revealed marked loss of cerebellar Purkinje cells, most severe in the vermis. Purkinje cell depletion was accompanied by Bergmann gliosis and empty baskets.\textsuperscript{21,35,36} In 2 individuals with CANVAS and an enriched phenotype including parkinsonism, marked neuronal depletion was observed in the substantia nigra and locus coeruleus.\textsuperscript{21,35} Sural nerve biopsies from individuals with RFC1 expansions diagnosed with CANVAS or chronic idiopathic axonal polyneuropathy demonstrated depletion of both large and small myelinated fibers with no evidence of active axonal regeneration.\textsuperscript{2,12,35}

**Major Phenotypic Characteristics of RFC1-Related Disease**

Since the discovery of expansions in RFC1, there has been considerable interest in defining the potential spectrum of associated phenotypes other than CANVAS.\textsuperscript{10-12,16,17,19,27,37-39} While this is a necessary and exciting step in the process of any evolving disease discovery, cautiousness and implementation of the same processes that were used to identify CANVAS remain important. Where the technology exists for objective measurement of clinical characteristics, these should be used in case identification. This particularly concerns the objective identification of bilateral vestibular hypofunction and somatosensory loss. Unfortunately, several studies have neglected to perform the relevant investigations and have instead relied on bedside assessments or incompletely described methodology.\textsuperscript{4,10-13,16} Key examples of this are the clinical vs video head impulse test (HIT) and the bedside vs electrophysiologic assessment of somatosensory function.\textsuperscript{40,41} The exception here may be an absence of cerebellar atrophy on MRI in the early stage of cerebellar disease where bedside examination findings may be more sensitive than MRI findings.\textsuperscript{40,41} Oculomotor abnormalities related to cerebellar impairment are generally present in cases of neurodegenerative CA and measurable by various oculographic modalities, even when structural changes in imaging are absent, and so, may function as an objective metric of cerebellar dysfunction. So far as the bedside examination of cerebellar function is concerned, assiduous examination for the presence of the clinical signs of cerebellar impairment is of utmost importance, understanding that those signs that do not alter in the presence of vestibular and/or somatosensory pathology are of greatest value.\textsuperscript{3,42} We did not aim to present an

**Figure 2 Visually Enhanced Vestibulo-Ocular Reflex (VVOR)**

- Head velocity
- Eye velocity

Left trace displays normal VVOR with gain \(\sim 1\) and few low amplitude corrective saccades. Right trace is that of an abnormal VVOR with a reduced gain and multiple salvos of high-amplitude corrective saccades. Red = head velocity; black = eye velocity.
exhaustive exposition of phenotypic features reported to date but rather to focus on those we believe are key to recognizing potential RFC1-related disease based on the more robust evidence in the literature.

Cerebellar Ataxia With Neuropathy and Vestibular Areflexia Syndrome

As discussed earlier, CANVAS presents with 3 key foci of pathology (Table 3):

Cerebellar Impairment
Individuals with CANVAS present a consistent pattern of cerebellar atrophy that preferentially involves the anterior and dorsal vermis (lobules VI, VIIa, and VIIb) and laterally, predominantly involves the crus I region. This is visible on MRI and has been confirmed on pathologic specimens. Affected individuals display a range of bedside cerebellar abnormalities including saccadic visual pursuit, gaze-evoked nystagmus, cerebellar dysarthria, and 4-limb ataxia.

Vestibular Hypofunction
The bilateral vestibular hypofunction in CANVAS is due to a vestibular neuropathy (Scarpa ganglionopathy) microscopically evident as a marked diminution in the number of vestibular ganglion cells. Clinically, this is most reliably and conveniently apparent as reduced vestibulo-ocular reflex gain on the video HIT. Involvement of the geniculate and trigeminal ganglia were also identified. Of clinical diagnostic relevance is that the cochlea nerve and end organ are unaffected in CANVAS, although common causes of hearing loss such as presbycusis may be a comorbidity.

Somatosensory Loss
The somatosensory impairment in CANVAS is that of a sensory neuronopathy. Neuropathology revealed marked neuronal loss in the dorsal root ganglia, and while this may manifest as reduced perception of proprioception and sensation on examination, detailed electrophysiologic studies are a far more robust means of identifying this component. The pattern of absent or reduced sensory nerve action potentials and cranial nerve involvement will aid in identifying the presence of sensory involvement and the differentiation of a neuropathy from a neuronopathy.

Other Features
The CANVAS phenotype is variably accompanied by autonomic dysfunction and chronic cough.
Pure Cerebellar Ataxia
To date, studies examining cohorts of individuals with ataxia for the presence of pathologic expansions of RFC1 demonstrate isolated pure CA is an infrequent finding.8,9 Pure CA is an ataxia of purely cerebellar origin, not involving other neurologic systems. Viewed in the context of the expanding phenotypes of other CAs, this is unsurprising, because even diseases such as SCA6, previously held up as the archetype of a pure cerebellar syndrome, have proven to be otherwise.46,47

Cerebellar Ataxia With Bilateral Vestibulopathy
The combination of cerebellar and bilateral peripheral vestibular impairment may have been the historical basis for the clinical definition of CANVAS,48 but it is not an invariable feature of RFC1-related disease. The true prevalence of this phenotype in RFC1-related disease is currently difficult to gauge because the recent retrospective studies, by their very nature, lack consistent data on formal vestibular function testing and may be under-recognizing the presence of peripheral vestibular hypofunction.4,13-16 Of interest Gisatulin et al.8 reported this phenotypic combination in one of the only studies to date where all participants have received objective vestibular assessment.

Somatosensory Loss
Our own cohort of individuals with CANVAS displays electrophysiologic evidence of a sensory neuronopathy, and this reflects the dorsal root ganglionopathy seen on post-mortem pathology.43 Because the phenotypes associated with RFC1-related disease expands beyond CANVAS, it is not surprising to find that the identification of somatosensory manifestations increases. These include sensorimotor impairment, as well as combinations of somatosensory impairment with CA or vestibulopathy.9,13-16 Somatosensory symptoms are a common presentation in neurologic practice with underlying pathologies including sensory neuropathies, many of which are diagnosed as idiopathic.49 Hence, the focus on RFC1-related isolated somatosensory impairment is well justified. As previously mentioned, accurate phenotyping is of great importance here and a case in point is 2 studies looking at the prevalence of isolated somatosensory abnormalities in individuals with biallelic pathologic expansions of RFC1. The 2 studies involved a combined total of 459 participants,10,12 none of whom received objective vestibular assessment.50 Given the known association of somatosensory loss and vestibulopathy, this methodology risks overestimating RFC1 as the cause of isolated somatosensory impairment. Similarly, the need for electrophysiologic investigation of any participant undergoing phenotyping for research on RFC1 is a necessity, while more detailed studies find utility in further characterization of any abnormality identified, for example, in defining a neuropathy vs a neuronopathy (ganglionopathy).45

Charcot-Marie-Tooth (CMT) disease may manifest as motor and/or sensory impairment. Lipponen et al.11 have suggested that RFC1 testing be included in the usual diagnostic workup of CMT; however, pending further clarification of phenotypes, it

---

**Figure 4** Comparison of Methods to Identify Repeat Expansions

<table>
<thead>
<tr>
<th>Flanking PCR</th>
<th>Repeat-primed PCR</th>
<th>Southern blot</th>
<th>Short-read sequencing</th>
<th>Long-read sequencing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advantages:</td>
<td>Advantages:</td>
<td>Advantages:</td>
<td>Advantages:</td>
<td>Advantages:</td>
</tr>
<tr>
<td>• Quick</td>
<td>• Quick</td>
<td>• Robust</td>
<td>• Can identify motif</td>
<td>• Can accurately size</td>
</tr>
<tr>
<td>• Can identify if there is a wild-type allele present</td>
<td>• Can identify motif-specific expansion</td>
<td>• Can accurately size large alleles</td>
<td>an expansion</td>
<td>large alleles</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Can identify the presence of an expansion</td>
<td>• Can determine entire sequence of repeat</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Can identify motif</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Can accurately size expansions</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Requires validation with PCR or Southern blotting</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Expensive</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Low per-read accuracy</td>
<td></td>
</tr>
</tbody>
</table>

Graphical representation of the different diagnostic methods for detecting repeat expansions. The samples analyzed for the flanking and repeat-primed PCR are taken from a patient with biallelic RFC1 expansions. ctrl = control; rpts = repeats.
is our opinion regarding the assertion that patients with CMT neuropathy were RFC1 positive that it is more likely that these patients have RFC1-related disease and not CMT.

Parkinsonian Disorders, Including Multiple Systems Atrophy

RFC1-related phenotypes may be similar to multiple systems atrophy (MSA), particularly the cerebellar type, in the early stage of MSA’s natural history. Several groups have interrogated MSA cohorts for RFC1 expansions with variable results. A clear phenotype-genotype relationship in MSA is challenging. One recent study found that only 62 percent of individuals who fulfill the clinical diagnostic criteria for MSA had their diagnosis upheld after autopsy. While biallelic RFC1 expansions were recently reported as a cause of MSA, the classification of cases as MSA was made purely using clinical diagnostic criteria. By contrast, Sullivan et al. did not find any evidence of biallelic RFC1 expansions in their cohort of 336 pathologically diagnosed individuals with MSA. Therefore, it seems more likely that phenotypic overlap in cases of RFC1-related disease with autonomic dysfunction and/or parkinsonism is the case. This may be an analogous situation to that of CMT discussed earlier and highlights the need to be cautious in attributing a previously described disease to that of CMT.

RFC1 Pathogenic Mechanisms

The pathogenic mechanism(s) underlying biallelic RFC1 expansions is currently unknown. There are multiple known mechanisms of repeat expansion pathogenesis including loss-of-function or gain-of-function, RNA-binding protein sequestration, and repeat-associated non-AUG (RAN) translation. RAN translation is a process whereby protein synthesis is initiated at the site of a repeat expansion in the absence of an AUG start codon. These RAN peptides can accumulate and form toxic aggregates. RAN translation has been associated with more than 10 repeat expansion disorders, but no evidence of RAN proteins has been observed in individuals with biallelic RFC1 expansions to date. Unlike other recessive repeat expansion disorders such as Friedreich ataxia, the RFC1 expansion does not appear to have a loss-of-function effect. In both peripheral tissues and postmortem brain samples, no reduction in RFC1 mRNA or protein was observed. In other conditions caused by intron repeat expansions, such as C9orf72 amyotrophic lateral sclerosis/frontotemporal dementia, altered RNA splicing results in intron retention and the accumulation of RAN proteins and RNA foci, both of which affect downstream pathways. To date, no abnormal splicing or intron retention has been consistently observed in individuals with biallelic RFC1 expansions, although a slight increase in the retention of intron 2 in pre-mRNA was reported in patient-derived lymphoblasts and muscle biopsies. While suggestive that abnormal pre-mRNA processing may play a role in the pathogenic mechanisms of RFC1 expansions, these studies were small (n = 2 and 7, respectively); therefore, additional studies with larger sample sizes are required to confirm these data.

RFC1 encodes the large subunit of replication factor C, a pentameric DNA polymerase accessory protein involved in DNA replication and repair. The RFC complex functions as a DNA-dependent ATPase that catalyzes the opening of the DNA-clamp protein proliferating cell nuclear antigen (PCNA), allowing it to encircle DNA and act as a scaffold to recruit DNA polymerases and other proteins involved in DNA replication and repair. RFC1 has been associated with various DNA damage response pathways including DNA damage recognition, recruitment of DNA repair machinery, base and nucleotide excision repair, mismatch repair, and promotion of cell survival. Of interest pathogenic variants in multiple genes involved in DNA repair pathways including tyrosyl-DNA phosphodiesterase 1, polyadenylate kinases 3’-phosphatase, X-ray repair cross-complementing 1, aprataxin, ataxia telangiectasia mutated, meiotic recombination 11 homolog A, and PCNA result in ataxia (Figure 3). This suggests that the cerebellum is particularly vulnerable to DNA damage. However, fibroblasts with biallelic RFC1 expansions do not show evidence of an impaired response to DNA damage. Furthermore, in a cohort of 100 White individuals with genetically confirmed biallelic RFC1 expansions, the incidence of neoplasm was not increased when compared with that in 100 randomly chosen age-matched and sex-matched individuals with ataxia, contrary to what might be expected if RFC1 function is impaired by the homozygous repeat expansion.

Molecular Testing for RFC1 Expansions

Based on the more robust elements of the RFC1 literature to date, we suggest testing for RFC1-related disease where the following phenotypes are accurately identified: CANVAS, late-onset pure CA, a sensory neuronopathy, combined CA and sensory neuronopathy or other peripheral sensory impairment, combined CA and a peripheral sensory motor neuropathy, any of these phenotypes in combination with autonomic dysfunction, parkinsonism, and/or a chronic cough. Where a family history of any of these phenotypes exists, the index of suspicion is raised. To date, it appears that RFC1-related disease generally has a very late onset and families with many affected individuals are relatively uncommon; hence, it may be difficult to discriminate between a dominant and recessive pattern, meaning that family history is an indicator but not exclusionary. The reader is directed to additional published guidance regarding clinical pathways for genetic testing of CANVAS and the genetic ataxias more broadly.

Molecular diagnosis of an RFC1 expansion requires a combination of a flanking PCR to test for the absence of a normal allele and repeat-primed PCR (RP-PCR) to identify the
presence of expanded alleles. Individuals with biallelic RFC1 expansions fail to generate a flanking PCR product but generate a characteristic sawtooth pattern by RP-PCR using a primer specific for the pathogenic repeat motif (Figure 4). Given the motif-specific nature of RP-PCR, assays targeting the AAGGG\(_{(n)}\) allele will fail to detect the alternate pathogenic repeat conformations such as the ACAGG\(_{(n)}\) allele. Accurate sizing of an expanded allele is challenging as it can range in size up to several thousand repeats, but size estimates can be generated by Southern blotting or long-read sequencing. Indeed, long-read sequencing is increasingly being used for the detection and size estimation of repeat expansions and has the potential to replace Southern blotting as the gold standard for sizing of expanded repeats.\(^{66}\) In addition, despite the intronic location, RFC1 expansions have been identified in both WGS and whole-exome sequencing data.\(^{3,38}\) Although not yet routinely incorporated into diagnostic pipelines, it is likely that in the near future, short-read diagnostic genome data will be interrogated using bioinformatic tools to identify expanded repeats,\(^{5}\) with subsequent orthogonal testing to validate the result and estimate the size of pathogenic alleles at specific loci performed, if appropriate. Currently, the availability of diagnostic testing for RFC1 expansion is very limited. Interrogation of the GTR resource (NCBI, 2-2022) identified a single diagnostic provider of RFC1 testing using flanking PCR and RP-PCR testing for the AAGGG\(_{(n)}\) pathogenic motif, although the authors were aware of multiple providers currently developing accredited tests.

Conclusions and Future Directions

Since the discovery in 1991 of the first expansion repeats, causing fragile X syndrome and spinal bulbar muscular atrophy, at least 50 repeat-mediated disorders have been described. Remarkably, more than half of these repeat expansions have been identified in the past 10 years, predominantly due to the advent of new sequencing technologies and associated bioinformatic analyses. Many of these expansions cause heterogeneous neurogenetic conditions that have proven difficult to both diagnose clinically and confirm with molecular genetic testing. For many of the recent discoveries, such as the repeats underlying familial cortical myoclonic tremor and CANVAS, the careful collection and deep phenotyping of affected family members has proven crucial to identification of the causal repeat expansion. The ability to achieve a molecular genetic diagnosis provides the opportunity to understand the natural history of a condition, and starting with CANVAS, we are just beginning a fascinating journey to understand RFC1-mediated disease. Additional deep phenotyping leveraging available objective methods will be essential in determining the true spectrum of phenotypes directly caused by RFC1 expansions. This process is also a necessary step in minimizing misdiagnoses, unnecessary or missed opportunities for prudent gene testing, and suboptimal management. Finally, we anticipate these studies will help to reveal the underlying disease mechanisms of pathogenic RFC1 expansions, which currently remain elusive. In an era of emerging therapies for repeat expansion disorders,\(^{67}\) understanding the mechanism will allow the identification of potential therapeutic targets and ultimately, the realization of safe and effective treatments.

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 March 1, 2022. Accepted in final form June 22, 2022. Submitted and externally peer reviewed. The handling editor was Massimo Pandolfo, MD, FAAN.

Appendix Authors

<table>
<thead>
<tr>
<th>Name</th>
<th>Location</th>
<th>Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kayli Davies, BSc Hons</td>
<td>Bruce Lefroy Centre for Genetic Health Research, Murdoch Children’s Research Institute; Department of Paediatrics, University of Melbourne, Parkville, Australia</td>
<td>Major role in the acquisition of data; study concept or design; and analysis or interpretation of data</td>
</tr>
<tr>
<td>David Szmulewicz, MBBS (Hons), PhD, FRACP</td>
<td>Balance Disorders &amp; Ataxia Service, Royal Victorian Eye &amp; Ear Hospital, East Melbourne; The Florey Institute of Neuroscience and Mental Health, Parkville, Australia</td>
<td>Drafting/revision of the article for content, including medical writing for content; analysis or interpretation of data</td>
</tr>
<tr>
<td>Louise A. Corben, PhD</td>
<td>Bruce Lefroy Centre for Genetic Health Research, Murdoch Children’s Research Institute; Department of Paediatrics, University of Melbourne, Parkville, Australia</td>
<td>Drafting/revision of the article for content, including medical writing for content; analysis or interpretation of data</td>
</tr>
<tr>
<td>Martin Delatycki, MD</td>
<td>Bruce Lefroy Centre for Genetic Health Research, Murdoch Children’s Research Institute; Department of Paediatrics, University of Melbourne, Parkville; Victorian Clinical Genetics Services, Melbourne, Australia</td>
<td>Drafting/revision of the article for content, including medical writing for content; analysis or interpretation of data</td>
</tr>
<tr>
<td>Paul J. Lockhart, PhD</td>
<td>Bruce Lefroy Centre for Genetic Health Research, Murdoch Children’s Research Institute; Department of Paediatrics, University of Melbourne, Parkville, Australia</td>
<td>Major role in the acquisition of data; study concept or design; and analysis or interpretation of data</td>
</tr>
</tbody>
</table>

References


