Supplement
Methods

Whole genome sequencing (WGS) was performed at Baylor Genetics Clinical Diagnostics Laboratory. Library preparation was performed using a PCR-free 550-bp insert size protocol by the KAPA Hyper Prep kit. Sequence analysis was performed on the Illumina NovaSeq platform for 150bp paired-end reads. Mean depth of coverage was 51x, and percent depth of coverage at 20x was 98.8%. Data analysis and interpretation were performed by the Baylor Genetics analytics pipeline. Data were aligned to the human reference genome build GRCh38 using the Illumina Dragen BioIT Platform. Variant calling was performed with the Illumina Dragen haplotype-based variant calling system and Illumina Dragen genome wide depth based CNV caller with custom modifications from Baylor Genetics. Trinucleotide repeat calling was performed using the Illumina Manta Structural Variant Caller. WGS data were analyzed for sequencing variants (SNV) and copy number variants (CNV) in nuclear genes and mitochondrial DNA (mtDNA) as well as short tandem repeats (STR). Expansion Hunter software was also used to analyze the WGS data on a research basis and did not identify any variants in known repeat expansion disorders other than the previously identified *RFC1* intronic expansion.

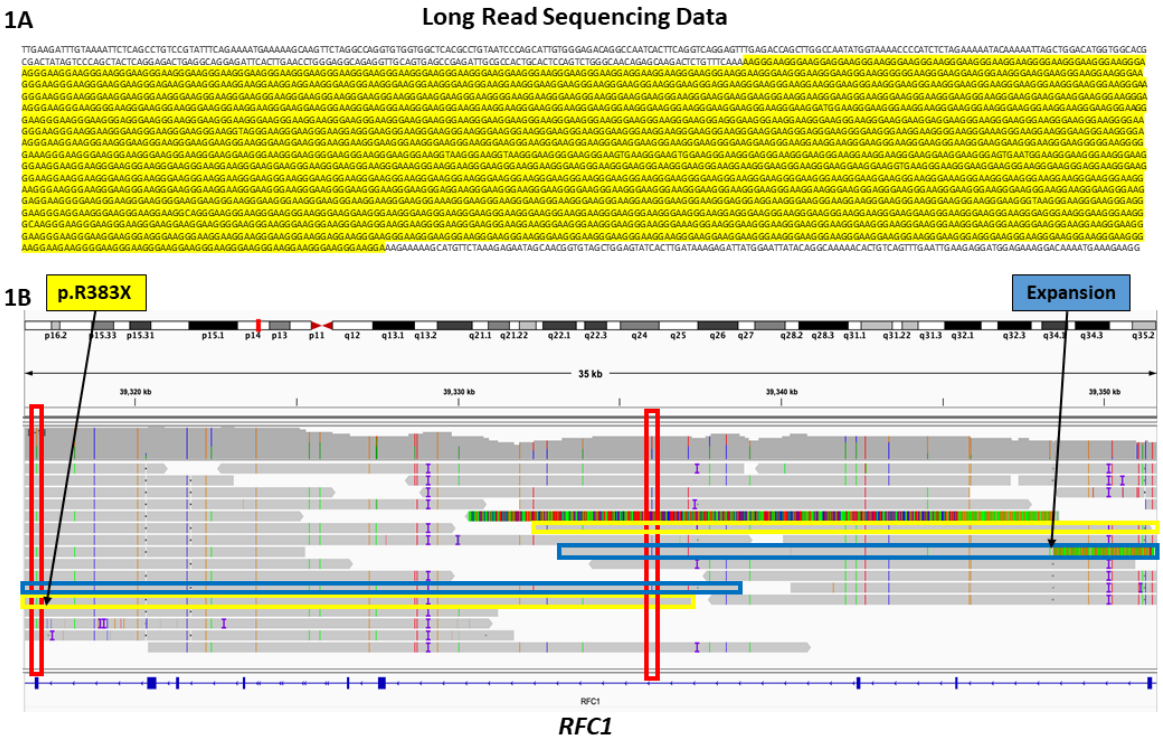
Only one other variant in a gene associated with ataxia was reported, a heterozygous splice site variant in *SPTBN2* (NM_006946.3:c.3777-3C>G), reported as a variant of uncertain significance. Variants in *SPTBN2* are associated with autosomal dominant (AD) spinocerebellar ataxia 5 (OMIM 600224) as well as autosomal recessive (AR) spinocerebellar ataxia-14 (OMIM 615386). The intronic *SPTBN2* variant is predicted to alter splicing and cause exon skipping, which would shift frame and result in a downstream premature stop codon and likely nonsense mediated decay. All evidence to date (gene pLI=0, none of the variants in OMIM associated with the AD form are loss-of-function variants) suggest that haploinsufficiency is not the mechanism for AD spinocerebellar ataxia 5, so this patient is likely a carrier for the recessive form of spinocerebellar ataxia-14. In addition, our patient's phenotype is not consistent with either the AD or AR forms of spinocerebellar ataxia. We did not identify any variants in *ELF2* or *PNPLA6*; variants in these genes have rarely been associated with the CANVAS phenotype.

Complementary DNA (cDNA) was synthesized from RNA obtained from peripheral blood from the proband and healthy control using SuperScript III.

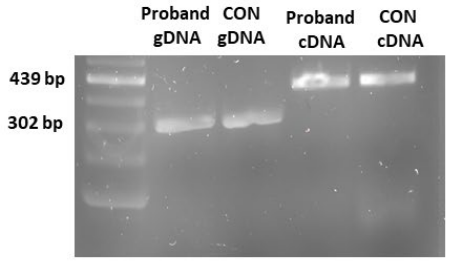
eFigure 1. We used PacBio long read sequencing (~20 kb reads, Pacific Biosciences, Menlo Park, California) to confirm the intronic expansion and to determine the intronic expansion and nonsense variant are in *trans.* We phased the variants using heterozygous variants in the interval

between the expansion and nonsense variant. 1A. Section of a single PacBio read spanning the repeat expansion in *RFC1*, demonstrating approximately 600 AAGGG repeats. 1B. Integrative Genome Viewer (IGV) screenshot of 35 Kb region of *RFC1*, containing both the p.R383X variant as well as the intronic repeat expansion. Phasing reveals the two variants are in *trans*, with the p.R383X variant on the allele represented with a yellow border, and the repeat expansion on the allele represented in a blue border. Variants in red boxes were used for phasing.

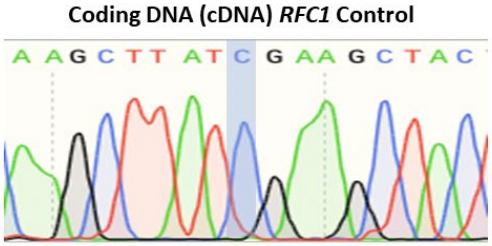
eFigure 2. 2A. Agarose gel electrophoresis of proband and control (CON) genomic DNA and cDNA (synthesized from RNA) bands including the region surrounding the p.R383X variant. Genomic DNA band size is 302 bp, and cDNA band size is 439 bp. 2B. Sequencing elecropherogram of control cDNA band demonstrates wild-type sequence (blue box).



2A.


Proband gDNA
CON gDNA
Proband cDNA
CON cDNA
439 bp
302 bp

2B.


Coding DNA (cDNA) RFC1 Control
A A G C T T A T C G A A G C T A C