RFC1 nonsense and frameshift variants cause CANVAS: clues for an unsolved pathophysiology.

Mehdi Benkirane, Dylan da Cunha, Cecilia Marelli, Lise Larrieu, Mathilde Renaud, Jessica Varilh, Morgane Pointaux, David Baux, Olivier Ardouin, Charles Vangoethem, et al.

▶ **To cite this version:**

Mehdi Benkirane, Dylan da Cunha, Cecilia Marelli, Lise Larrieu, Mathilde Renaud, et al.. RFC1 nonsense and frameshift variants cause CANVAS: clues for an unsolved pathophysiology.. Brain - A Journal of Neurology , 2022, ⟨10.1093/brain/awac280⟩. ⟨hal-03750598⟩

HAL Id: hal-03750598

https://hal.science/hal-03750598v1

Submitted on 24 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HAL Authorization

RFC1 nonsense and frameshift variants cause CANVAS: clues for an unsolved pathophysiology

Mehdi Benkirane,1,2,† Dylan Da Cunha,2,† Cecilia Marelli,3,4 Lise Larrieu,1 Mathilde Renaud,5 Jessica Varilh,2 Morgane Pointaux,1 David Baux,1 Olivier Ardouin,1 Charles Vangoethem,1 Magali Taulan,2 Benjamin Daumas Duport,6 Anne Bergougnoux,1,2 Anne-Gaelle Corbillé,7 Mireille Cossée,1,2 Raul Juntas Morales,3 Sylvie Tuffery-Giraud,2 Michel Koenig,1,2 Bertrand Isidor8 and Marie-Claire Vincent1,2

†These authors contributed equally to this work.

1 Department of Molecular Genetics, Institut Universitaire de Recherche Clinique (IURC), Montpellier Hospital, Montpellier, France

2 Genetics and Pathophysiology of NeuroMuscular Disorders, PhyMedExp Research Unit, CNRS, INSERM, University of Montpellier, Montpellier, France

3 Department of Neurology, Montpellier Hospital, Montpellier, France

4 Molecular Mechanisms of Neurodegenerative Dementia (MMDN), EPHE University of Montpellier, INSERM, Montpellier, France

5 Department of Medical Genetics, Nancy Hospital, Nancy, France

6 Department of Radiology, Nantes Hospital, Nantes, France

7 Department of Neurology, Nantes Hospital, Nantes, France

8 Department of Medical Genetics, Nantes Hospital, Nantes, France

Correspondence to:
Professor Michel Koenig, MD, PhD Institut Universitaire de Recherche Clinique 641 Avenue du Doyen Gaston Giraud 34093 Montpellier Cedex 5, France E-mail: michel.koenig@inserm.fr

Keywords:

CANVAS; *RFC1*; cerebellar ataxia; truncating variant; repeat expansion

Cerebellar ataxia, neuropathy and vestibular areflexia syndrome (CANVAS) is an inherited late-onset neurological disease caused by bi-allelic AAGGG pentanucleotide expansions within intron 2 of *RFC1*. Despite extensive studies, the pathophysiological mechanism of these intronic expansions remains elusive.

We screened by clinical exome sequencing two unrelated patients presenting with late-onset ataxia. A repeat-primer polymerase chain reaction was used for *RFC1* AAGGG intronic expansion identification. *RFC1* mRNA expression was assessed by quantitative reverse transcription–polymerase chain reaction.

We identified the first two CANVAS affected patients who are compound heterozygous for *RFC1* truncating variants (p.Arg388* and c.575delA, respectively) and a pathological AAGGG expansion. *RFC1* expression studies in whole blood showed a significant reduction of *RFC1* mRNA for both patients compared to three patients with bi-allelic *RFC1* expansions.

In conclusion, this observation provides clues that suggest bi-allelic *RFC1* conditional loss-of-function as the cause of the disease.

Introduction

Cerebellar ataxia, neuropathy and vestibular areflexia syndrome (CANVAS) is a recessive neurodegenerative disease, estimated to be one of the most common recessive ataxia.1Clinical features associated with CANVAS include late-onset cerebellar ataxia (80%), length-independent sensory neuronopathy (100%) and bilateral vestibular impairment (53%).2Less frequently, autonomic dysfunction, parkinsonism and pyramidal irritation are noticed, indicating overlap with multiple system atrophy.3Recently, bi-allelic expansions within an AluSx3 element in intron 2 of *RFC1* have been identified as the cause of CANVAS.1,4To date, only bi-allelic expansions of the AAGGG pentanucleotide motif or an East Asian ACAGG motif have been reported to cause the disease, although other pathogenic motifs arrangement exist.5,6CANVAS pathophysiology has remained elusive so far, since previous studies failed to demonstrate direct loss-of-function of *RFC1*.1Herein we report two patients from unrelated families harbouring compound heterozygous pathogenic AAGGG expansion and *RFC1* truncating mutation (p.Arg388* and c.575delA, respectively). Studies of *RFC1* messenger RNA expression in whole blood of these two patients compared to bi-allelic expansion patients provide pathophysiological clues for this enigmatic disease.

Materials and methods

Patient and clinical information collection

The proband of Family 1 (Patient F1-II.4) was part of a large ataxia cohort of 448 families, including 366 families previously reported.7The proband of Family 2 (Patient F2.II.1) was referred for *RFC1* pathogenic expansion screening due to suspected CANVAS diagnosis. Both patients were analysed by clinical exome sequencing after exclusion of expansions causing spinocerebellar ataxias 1, −2, −3, −6, −7, −17, dentato-rubro-pallido-luysian atrophy and Friedreich's ataxia. Clinical data were collected by B.I. and A.-G.C. for Patient F1-II.4, and C.M. and R.J.M. for Patient F2-II.1, respectively. All individuals signed a written consent prior to genetic analysis. All procedures were conducted according to the guidelines of the Declaration of Helsinki.

Clinical exome sequencing and data analysis

Clinical exome sequencing was performed using the TruSight One expanded kit (www.illumina.com/trusightone) that targets ~6700 genes, most of them involved in human inherited diseases. Sequencing was performed using Illumina NextSeq technology with 150-bp paired-end reads (Montpellier Hospital next- generation sequencing platform). Sequence alignment, variant calling was performed following the Broad Institute's Genome Analysis Toolkit (GATK4) best practices, as described in a previous study.7Guidelines from the American College of Medical Genetics were used for clinical sequence interpretation.8BAM files were viewed with Integrative Genomics Viewer software (v.2.7.0 Broad Institute). Probands and available relatives from each family were tested for mutation segregation by Sanger sequencing.

Repeat-primer PCR and long-range PCR analysis

Repeat-primer PCRs (RP-PCRs) were designed for each of three pentanucleotide repeat units (AAAAG/AAAGG/AAGGG). Fragment length analyses were performed on an ABI3500xl/Dx DNA Analyzer (Applied Biosystems®) and visualized through GeneMapper® v.4.0 (Applied Biosystems®). Presence of normal-sized or expanded repeat alleles was confirmed by standard and long- range PCR (LR-PCR), using Phusion® High-Fidelity PCR Master Mix (New England Biolabs). Sanger sequencing was performed on expanded alleles to determine the repeat motif sequence. Primers, reagents and methods were designed as described in Cortese *et al.*1

Total RNA Isolation and *RCF1* transcripts analysis

Total RNA was purified after collection of peripheral whole blood using 2.5 ml of the PAXgene Blood RNA kit (Qiagen, Courtaboeuf, France) allowing intracellular RNA stabilization for quantification. First-strand cDNAs were synthesized from 300 ng of total RNA using random primers and the SuperScript III First Strand cDNA Synthesis Kit (Invitrogen). For quantitative reverse transcription PCR (RT-PCR), 1 µl of 1:6 diluted cDNA was amplified using the LightCycler 480 SYBR Green I Master Kit in a final volume of 5 µl containing 0.5 µM of primers. Three different primer pairs spanning *RFC1* exons 2/3 (F1-R1) and exons 7/8 (F2-R2), previously described,1and exons 20/21 (F3-R3) were used for *RFC1* amplification. Each sample was assayed in triplicate. Amplified transcripts were quantified using the comparative CT method normalized to *GAPDH* (glyceraldehyde-3- phosphate dehydrogenase) and *RPLP0* (ribosomal protein, large, P0) (QuantiTect Primer Assay, Qiagen) as reference housekeeping transcripts, and presented as normalized fold expression change (2ΔΔCT). The differences in the expression of *RFC1* transcripts between biological conditions were tested using the non-parametric Kruskal– Wallis statistical test with the GraphPad Prism 9 software (version 4.0). Technical replicate data for the three *RFC1* exons were pooled for each condition (controls n=36; E/E n=27; Patient F2-II.1 n=9; Patient F1-II.4 n=9). A P-value ≤0.05 was considered statistically significant. Inclusion level of mutated *RFC1* exons 6 and 10 was assessed by PCR amplification (35 cycles) using primers spanning exons 5/7 (F4-R4) and exons 9/11 (F5-R5), respectively. PCR products were analysed by agarose gel electrophoresis and Sanger sequencing (a list of primers in provided in Supplementary Table 1). Semi-quantitative allelic contribution to *RFC1* transcript expression was estimated based on calibrated nucleotide peak height from Sanger sequence electropherograms. Electropherogram peak heights were adjusted with nucleotides flanking the mutations. Allele peak heights were estimated by comparison with the corresponding wild-type nucleotide of genomic DNA and cDNA from a control individual (set at the 100% level) and with wild-type and mutant nucleotides of genomic DNA from the patients (both set at the 50% level).

Ethics declaration

The study was approved by the local ethics committee. All individuals signed a written consent before genetic analysis.

Data availability

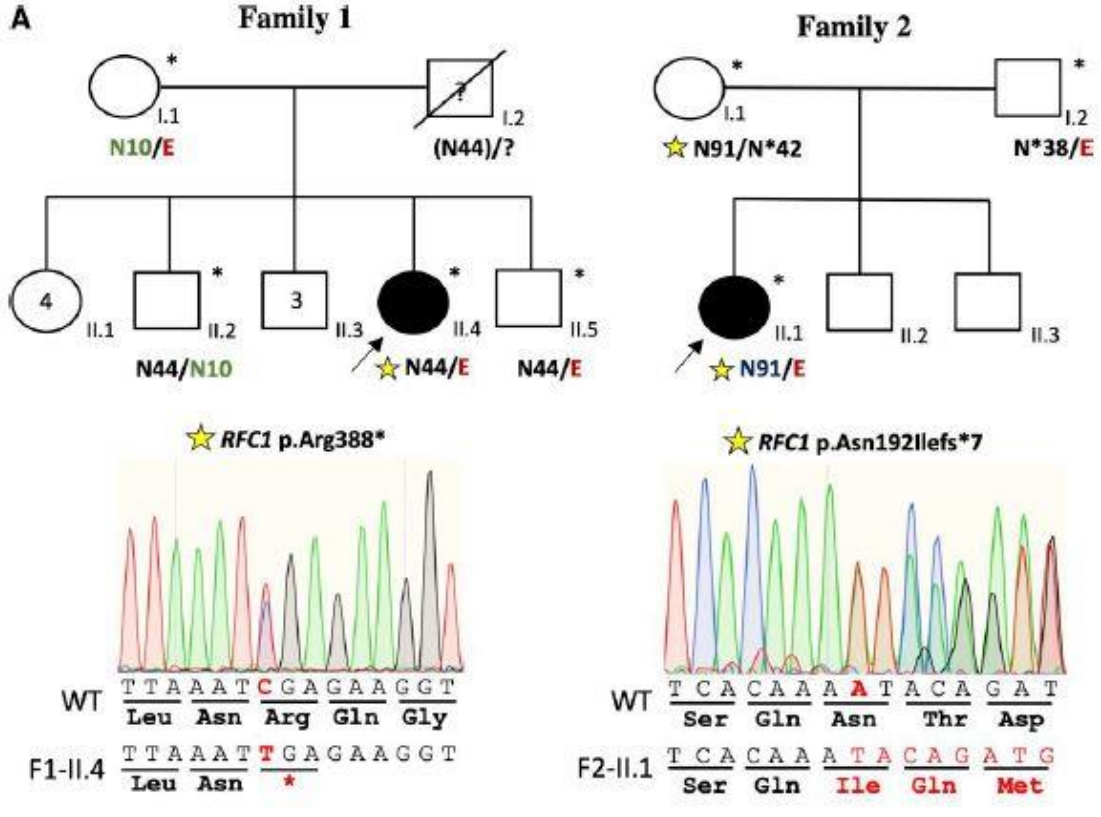
Readers can freely access all software used for this study by following the most recent version links mentioned in the 'Materials and methods' section. Patients' data can be made available upon reasonable request.

Results

Clinical findings

The proband of Family 1 (Patient F1-II:4) is a 68-year-old female born from unrelated parents who presented with progressive cerebellar and sensory ataxia at 45 years. Cerebellar signs included dysmetria, understandable dysarthria and gait imbalance leading to walking disability without support at 59 years. Neurological examination confirmed lower limbs hypoesthesia, length- dependent paraesthesia and loss of vibration sense. Deep tendon reflexes were brisk on the upper limbs and absent on the lower limbs. She later exhibited autonomic dysfunction with bladder sphincter complains, requiring permanent protection wearing. Vestibulo–ocular reflexes were clinically preserved on last examination (at 66 years).

Similarly, a 61-year-old female (Patient F2-II:1) presented with length-dependent sensory impairment, including painful hyperesthesia and hypopallesthesia of the lower limbs since the age of 46 years, associated with chronic cough. Mild gait ataxia, brisk lower limbs reflexes and bilateral loss of vestibulo–ocular reflexes were noticed on first examination. Disease progression led to upper limb sensory symptoms (49 years), dysphagia, dysarthria (57 years) and walking disability without support at 60 years. In both cases, electroneuromyography studies revealed normal motor nerve conductions and reduced nerves sensory amplitudes from the upper and lower limbs, consistent with sensory neuronopathy and dorsal root ganglia degeneration. Cerebral MRI showed global cerebellar atrophy for both patients, and non-specific diffuse white matter hyper intensities only for Patient F1-II.4 (Fig. 1B).



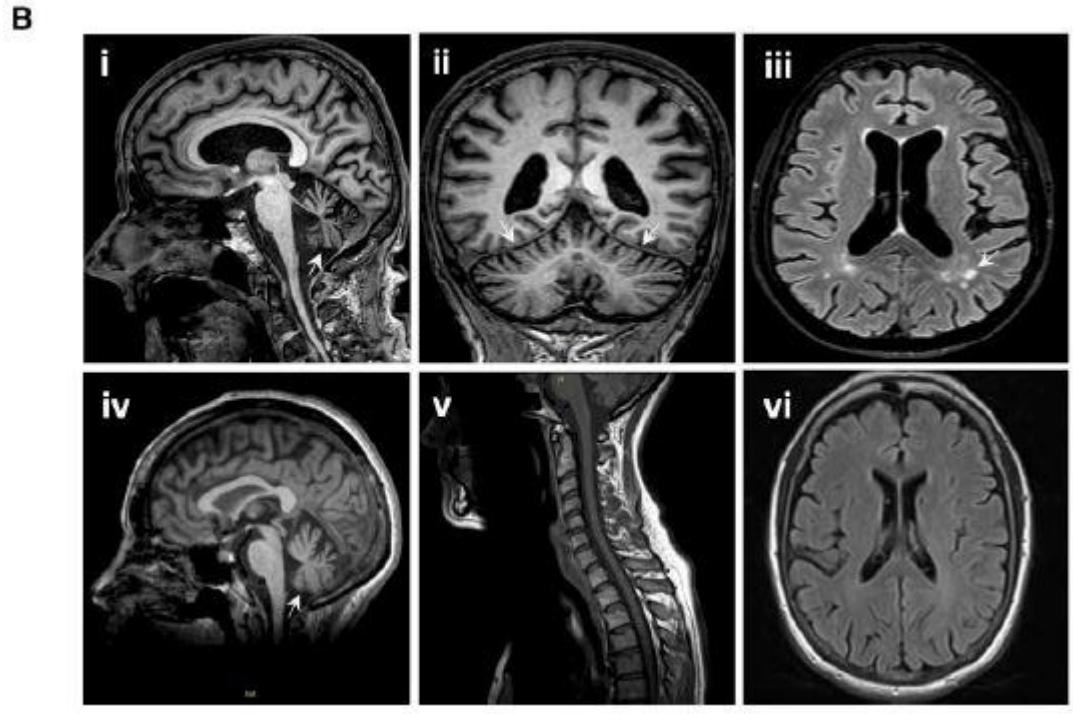
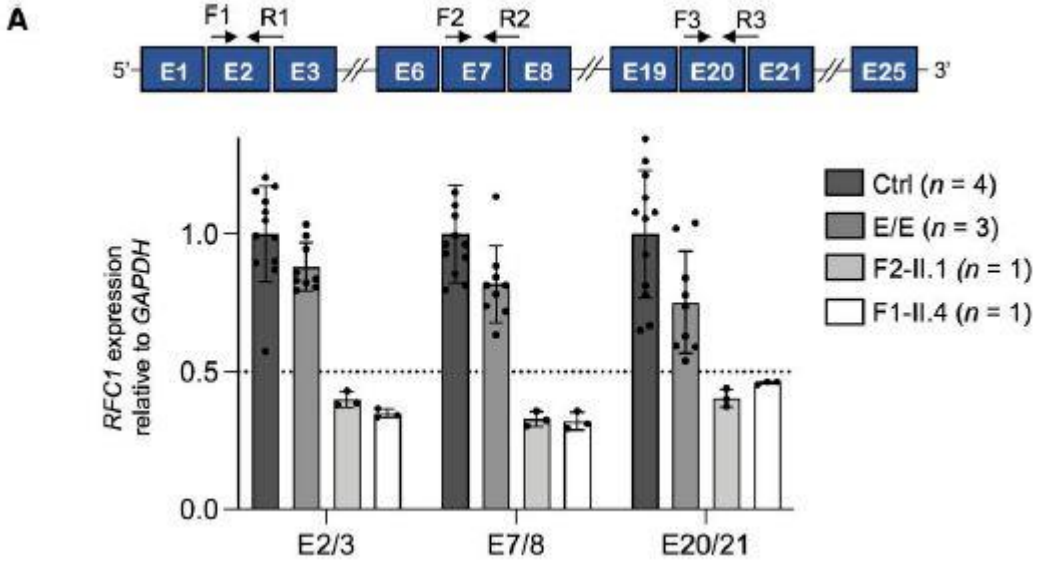


Figure 1 Pedigrees and cerebral MRI of compound heterozygous patients. (**A**) Pedigree of Families 1 and 2 showing segregation of truncating *RFC1* variants (p.Arg388* and c.575delA, yellow star) and intronic AAGGG *RFC1* expansion (E). N indicates normal AAAAG allele followed by the number of repeats. N* indicates non-pathogenic expanded alleles associating different motifs followed by the number of repeats. (**B**) Cerebral and spinal cord MRI of both probands. Patient F1-II.4 (**i**–**iii**): sagittal (**i**) and coronal (**ii**) T1-weighted images of cerebral sections showing marked vermian and hemispheric cerebellar atrophy (arrows), respectively. Axial FLAIR image (**iii**) showing white matter hyper intensities (arrow). Patient F2-II.1 (**iv**–**vi**): sagittal T1-weighted images of cerebral sections (**iv**) and spinal cord section (**v**) showing moderate global cerebellar atrophy (arrow) and normal image, respectively. Normal axial FLAIR image (**vi**) of cerebral sections.



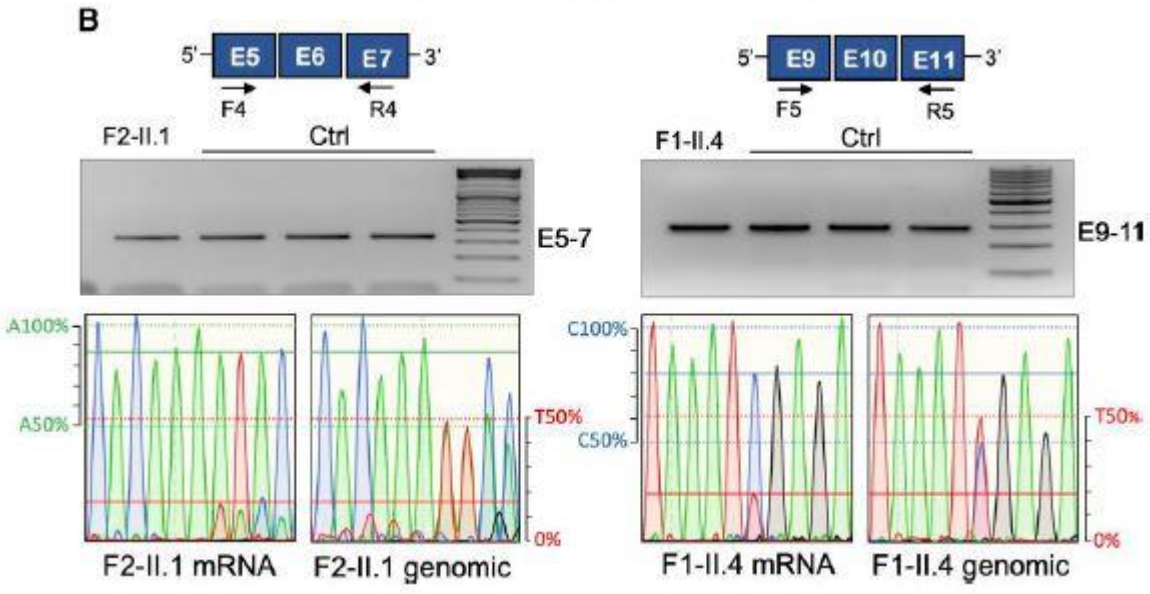


Figure 2 *RFC1* mRNA transcript analysis. (**A**) Schematic representation of *RFC1* transcript (NM_002913.5) and mapping of the primers (black arrows) used for *RFC1* mRNA assessment by quantitative RT-PCR (F1-R1, F2-R2 and F3-R3) in controls (*n*=4), E/E affected individuals (*n*=3) and both reported patients (F2-II.1 and F1-II.4). *GAPDH* was used as reference housekeeping gene for normalization. Data of technical replicates for each tested individual are reported (black dots). The bar graphs show the mean±standard deviation. (**B**) Agarose gel electrophoresis of RT-PCR products of exons 5/7 (*left*) and of exons 9/11 (*right*) in controls (Ctrl *n*=3) and Patients F2-II.1 or F1-II.4 with the corresponding sequences of *RFC1* exons 6 and 10 amplified from mRNA and genomic DNA (primers F4-R4, F5-R5). Electropherograms were normalized according to peak heights of nucleotides preceding the point mutations in exons 6 and 10. Contribution of both expanded and truncated alleles in RT-PCR product was calibrated against homozygotes (shown in Supplementary Fig. 2) and against the corresponding genomic DNA (50%). Scales used for level estimation of the mutant alleles are indicated.

Genetic investigations

A heterozygous *RFC1* nonsense mutation c.1162C>T (p.Arg388*) was identified by clinical exome sequencing in the proband of Family 1 (Patient F1-II.4) and was absent in her mother and two unaffected brothers (the father being deceased). An AAGGG expansion was subsequently found by RP-PCR in the proband, her mother and one unaffected brother, while a large normal (AAAAG)44 repeat allele was identified in all three siblings but not in the mother (Fig. 1Aand Supplementary Fig. 2). Flanking single-nucleotide polymorphism analysis confirmed the *de novo* occurrence of the c.1162C>T variant, presumably on the father's allele, in *cis* with the large normal (AAAAG)44 allele (Supplementary Table 2and Supplementary Fig. 3). Similarly, *RFC1* screening of the second patient (Patient F2-II.1) revealed a heterozygous AAGGG expansion inherited from her father, and a maternally inherited c.575delA (p.Asn192Ilefs*7) frameshift mutation found by clinical exome sequencing (Fig. 1A).

RFC1 transcript analyses

RFC1 mRNA quantification in whole blood of both probands, three patients with bi-allelic RFC1 AAGGG expansion (E/E) and four control individuals was assessed by quantitative RT-PCR. Using GAPDH as reference housekeeping gene, a marked reduction (>50%) of RFC1 transcripts level was observed for both compound heterozygous probands F2-II.1 and F1-II.4, while E/E patients did not show a significant change in RFC1

transcript expression compared to controls (Fig. 2A). These results were comparable for the three primer sets spanning exons 2–3, exons 7–8 and exons 20–21, and were confirmed with a second reference gene (RPLP0) (Supplementary Fig. 1). Statistical analyses of all data-points (shown separate for each individual in Supplementary Fig. 1) indicate that *RFC1* mRNA reduction in proband F1-II.4 was highly significant compared to controls ($P<0.01$) using both reference housekeeping genes for normalization. For proband F2-II.1, this reduction was also highly significant when normalized with *GAPDH* ($P=0.003$) but only suggestive ($P=0.142$, ns) when *RPLP0* was used as reference gene (Supplementary Fig. 1). *RFC1* transcripts levels were also significantly reduced ($P<0.05$) in probands F1-II.4 and F2-II.1 compared to E/E patients. We next sought to estimate the contribution of each allele to this reduction by visualization of the truncating allele on sequences of the corresponding complementary DNA fragments. Semi-quantitative estimation of the respective transcripts by comparison with genomic DNA sequencing peak heights (Supplementary Fig. 2) indicates that the c.1162C>T and c.575delA alleles account for ≈20 and 15% of total RT-PCR products, respectively, reflecting a significant expression reduction of both truncating alleles down to 10% (Fig. 2B). Transcript level of the AAGGG expanded allele was deduced to be on the normal range for both patients (Patients F1-II.4 and F2-II.1) compared to controls and E/E patients. RT-PCR analysis of fragments, spanning exons 5 to 7 and exons 9 to 11, did not show abnormal splicing (Fig. 2Band Supplementary Fig. 4).

Discussion

We expand the mutational spectrum of CANVAS causing variants by reporting the first two patients who are compound heterozygous for the pathogenic AAGGG expansion and an *RFC1* mutation (p.Arg388* and c.575delA, respectively) and who present with a late-onset form of cerebellar ataxia consistent with CANVAS diagnosis. Age of onset (45 and 46 years) and clinical presentation appear unremarkable from that of E/E affected individuals (average 54±9 years, range 35–73 years).1,2,9Quantification of *RFC1* transcripts by quantitative RT-PCR in whole blood showed a marked reduction of expression of the p.Arg388* and c.575delA alleles, reflecting efficient nonsense-mediated mRNA decay of the truncating alleles. E/E affected individuals did not show significant reduction mRNA expression, in accordance with previous results obtained by Cortese and colleagues1in several other tissues.

RFC1 loss-of-function evidence?

RFC1 encodes the large subunit of Replication Factor C1, a heteropentameric AAA ATPase that acts as PCNA clamp-loader, particularly involved in mismatch repair, nucleotide excision repair and translesion synthesis.10,11Interestingly, numbers of neurodegenerative diseases leading to ataxia and neuropathy are caused by genes involved in DNA repair pathways12,13(e.g. *ATM*, *MRE11*, *APTX*, *PNKP*, *ERCC6* and *ERCC8*), reflecting cerebellar and sensory neurons intolerance to DNA damage. Pathophysiology of CANVAS remains elusive since no detectable impact was observed on *RCF1* gene expression or splicing in peripheral tissues or post-mortem brain samples (including cortex and cerebellum).1In addition, no RNA foci of either the sense (AAGGG) or antisense (CCCUU) repeated unit could be detected in any of the examined tissues.12,14. Identification of truncating *RFC1* variants leading to CANVAS is suggestive of a bi-allelic loss-of-function of *RFC1* as the cause of the disease. The normal *RFC1* mRNA transcript level found in leucocytes and other tissues of E/E affected individuals suggests an incomplete impact of the intronic *RFC1* AAGGG expansion on *RFC1* transcriptional regulation, by contrast with similar bi-allelic non-coding expansion diseases such as Friedreich's ataxia.15Rather, the AAGGG intronic expansion may only conditionally interfere with *RFC1* expression in stress conditions, not seen in native leucocytes or in post-mortem tissues (where fragile cells have already disappeared). Stress conditions may relate to DNA repair at the *RFC1* locus and may account for the specific pathogenicity of AAGGG expansions, as opposed to AAAGG and AAAAG expansions, possibly in relation with the GC content of the expanded motifs.16Alternatively, an RNA toxic pathogenic mechanism for the AAGGG expansion cannot be formally excluded, as was suggested by foci detection in SHY-H5 cells transfected with (AAGGG)54 oligonucleotides, and by the identification of intron 2 retention in *RFC1* transcripts across different tissues of

patients with bi-allelic expansion.1In this case, one normal allele, absent from our compound heterozygous probands, would be sufficient to balance the recessive toxic gain of function of the expanded allele. Identification of additional CANVAS patients with non-expanded mutations is needed to unravel the puzzling pathophysiology of this intriguing disease.

Acknowledgements

We sincerely thank the patients and families for participation to this study. We thank Pr. Vincent Procaccio for contribution to the etiological investigations.

Funding

This work was supported in part by the French patients' association 'Connaître les Syndromes Cérébelleux' (CSC). D.D.C. is supported by a fellowship from the French Association against Myopathies (AFM-Telethon).

Competing interests

The authors report no competing interests

Supplementary material

Supplementary materialis available at *Brain* online.

References

1. 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;51:649–658.

2. Cortese A, Tozza S, Yau WY, *et al.* Cerebellar ataxia, neuropathy, vestibular areflexia syndrome due to RFC1 repeat expansion. *Brain*. 2020;143:489–490.

3. Traschütz A, Cortese A, Reich S, *et al.* Natural history, phenotypic spectrum, and discriminative features of multisystemic RFC1 disease. *Neurology*. 2021; 96:e1369–e1382.

4. 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.

5. Scriba CK, Beecroft SJ, Clayton JS, *et al.* A novel RFC1 repeat motif (ACAGG) in two Asia-pacific CANVAS families. *Brain*. 2020;143: 2904–2910.

6. Beecroft SJ, Cortese A, Sullivan R, *et al.* A Maori-specific RFC1 pathogenic repeat configuration in CANVAS, likely due to a founder allele. *Brain*. 2020;143:2673–2680.

7. Benkirane M, Marelli C, Guissart C, *et al.* High rate of hypomorphic variants as the cause of inherited ataxia and related diseases: Study of a cohort of 366 families. *Genet Med*. 2021;23: 2160–2170.

8. Richards S, Aziz N, Bale S, *et al.* Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405–424.

9. Huin V, Coarelli G, Guemy C, *et al.* Motor neuron pathology in CANVAS due to RFC1 expansions. *Brain*. 2021;145:2121–2132.

10. Arbel M, Choudhary K, Tfilin O, Kupiec M. PCNA loaders and unloaders—One ring that rules them all. *Genes (Basel)*. 2021;12: 1812.

11. Umar A, Buermeyer AB, Simon JA, *et al.* Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell*. 1996;87:65–73.

12. Depienne C, Mandel JL. 30 Years of repeat expansion disorders: What have we learned and what are the remaining challenges? *Am J Hum Genet*. 2021;108:764–785.

13. Renaud M, Tranchant C, Koenig M, Anheim M. Autosomal recessive cerebellar ataxias with elevated alpha-fetoprotein: Uncommon diseases, common biomarker. *Mov Disord*. 2020;35: 2139–2149.

14. Malik I, Kelley CP, Wang ET, Todd PK. Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat Rev Mol Cell Biol*. 2021;22:589–607.

15. Li Y, Lu Y, Polak U, *et al.* Expanded GAA repeats impede transcription elongation through the FXN gene and induce transcriptional silencing that is restricted to the FXN locus. *Hum Mol Genet*. 2015;24:6932–6943.

16. Sznajder ŁJ, Thomas JD, Carrell EM, *et al.* Intron retention induced by microsatellite expansions as a disease biomarker. *Proc Natl Acad Sci U S A*. 2018;115:4234–4239.