



Analysis and occurrence of biallelic pathogenic repeat expansions in *RFC1* in a German cohort of patients with a main clinical phenotype of motor neuron disease

Annalisa Schaub^{1,2} · Hannes Erdmann^{1,2} · Veronika Scholz¹ · Manuela Timmer³ · Isabell Cordts^{4,5} · Rene Günther⁶ · Peter Reilich² · Angela Abicht^{1,2} · Florian Schöberl²

Received: 1 May 2024 / Revised: 9 June 2024 / Accepted: 11 June 2024 / Published online: 25 June 2024

© The Author(s) 2024

Abstract

Biallelic pathogenic repeat expansions in *RFC1* were recently identified as molecular origin of cerebellar ataxia, neuropathy, vestibular areflexia syndrome (CANVAS) as well as of one of the most common causes of adult-onset ataxia. In the meantime, the phenotypic spectrum has expanded massively and now includes mimics of multiple system atrophy or parkinsonism. After identifying a patient with a clinical diagnosis of amyotrophic lateral sclerosis (ALS) as a carrier of biallelic pathogenic repeat expansions in *RFC1*, we studied a cohort of 106 additional patients with a clinical main phenotype of motor neuron disease (MND) to analyze whether such repeat expansions are more common in MND patients. Indeed, two additional MND patients (one also with ALS and one with primary lateral sclerosis/PLS) have been identified as carrier of biallelic pathogenic repeat expansions in *RFC1* in the absence of another genetic alteration explaining the phenotype, suggesting motor neuron disease as another extreme phenotype of *RFC1* spectrum disorder. Therefore, MND might belong to the expanding phenotypic spectrum of pathogenic *RFC1* repeat expansions, particularly in those MND patients with additional features such as sensory and/or autonomic neuropathy, vestibular deficits, or cerebellar signs. By systematically analyzing the *RFC1* repeat array using Oxford nanopore technology long-read sequencing, our study highlights the high intra- and interallelic heterogeneity of this locus and allows the identification of the novel repeat motif ‘ACAAG’.

Keywords Repeat analysis · Oxford nanopore sequencing · Amyotrophic lateral sclerosis · RFC1 · MND · Motor neuron disease

Annalisa Schaub and Hannes Erdmann contributed equally.

✉ Florian Schöberl
florian.schoeberl@med.uni-muenchen.de

- ¹ Medical Genetics Center, Munich, Germany
- ² Department of Neurology With Friedrich-Baur-Institute, Klinikum Der Universität, Ludwig-Maximilians-University, Marchioninistr. 15, 81377 Munich, Germany
- ³ Gemeinschaftspraxis Für Humangenetik Dresden, Medizinische Genetik, Dresden, Germany
- ⁴ Department of Neurology, Klinikum Rechts Der Isar, Technical University of Munich, Munich, Germany
- ⁵ Department of Neuroscience, Mayo Clinic, Jacksonville, FL, USA
- ⁶ Department of Neurology, Universitätsklinikum Carl Gustav Carus Dresden, Dresden, Germany

Abbreviations

| | |
|--------|--|
| ALS | Amyotrophic lateral sclerosis |
| Exp | Expanded |
| Var | Variable |
| MND | Motoneuron disease |
| NGS | Next-generation sequencing |
| ONT | Oxford nanopore technologies |
| ROI | Region of interest |
| CANVAS | Cerebellar ataxia, sensory neuropathy, and vestibular areflexia syndrome |
| MSA | Multiple system atrophy |
| INDEL | Insertion and deletion |
| PLS | Primary lateral sclerosis |
| PMA | Progressive muscle atrophy |
| HSP | Hereditary spastic paraparesis |

Introduction

In 2019, biallelic pathogenic repeat expansions in intron 2 of the *replication factor C subunit 1* gene (*RFC1*) have been identified as a common genetic cause of cerebellar ataxia, neuropathy, vestibular areflexia syndrome (CANVAS) and were shown to be a major cause of hereditary late-onset ataxia [1, 2]. Unlike other repeat expansion disorders, *RFC1* spectrum disorder does not only require the expansion of a wild-type motif but also its substitution for a pathogenic motif (most commonly ‘AAGGG’). The *RFC1* locus is highly polymorphic. To date, at least 11 motifs are known with 6 of which are likely pathogenic when largely expanded [3–7].

The phenotypic spectrum associated with these repeat expansions is broad and includes chronic cough, sensory neuropathy, cerebellar ataxia, bilateral vestibular failure, and autonomic dysfunctions, among others [8, 9]. Interestingly, biallelic repeat expansions in *RFC1* have also been associated with several other distinct neurodegenerative disorders, such as Parkinson’s disease and atypical parkinsonism such as multiple system atrophy (MSA) [10–12]. Recently, the clinical phenotypic spectrum of *RFC1*-related disorders has been further extended by the report of upper and lower motor neuron involvement observed clinically as well as in post-mortem neuropathological examinations [13, 14]. Furthermore, a case of ALS and additional sensory neuropathy and presbyvestibulopathy in the presence of biallelic pathogenic *RFC1* motifs has been reported [15].

This raises the question of whether *RFC1* should be added to the list of MND-related genes and considered in the genetic diagnostic process. This is particularly relevant, considering that the etiology of sporadic MND such as amyotrophic lateral sclerosis (ALS) and primary lateral sclerosis (PLS) remains unexplained in up to ~90% of cases, and biallelic repeat expansions in *RFC1*, an autosomal-recessive disease, may potentially offer an additional genetic explanation in at least a subset of those patients [16–18].

Therefore, in this study, we conducted a systematic screening for *RFC1* expansions in a cohort of 107 patients with MND primarily composed of ALS patients employing CRISPR/Cas9-targeted Oxford nanopore technology (ONT) long-read sequencing [4, 19]. This method allowed not only an in-depth analysis of *RFC1*, evaluating repeat motifs and accurate repeat lengths in a large cohort, but also a combined assessment of the important ALS-related *C9orf72* repeat array. In addition, we applied a comprehensive high-throughput sequencing gene panel of 380 MND-related genes to all individuals to exclude additional relevant and so far known underlying genetic alterations.

Materials and methods

Patient cohort

All patients received diagnostic testing as ordered by their attending physicians (e.g., neurologist, geneticist). Informed consent was signed and obtained in all the cases (see ethical standards). Patients were informed about the findings and their significance for relatives. Genetic counseling was recommended and offered in all these cases.

Extraction of genomic DNA (gDNA)

Genomic DNA (gDNA) of 107 patients with a clinical phenotype of MND was obtained from total peripheral EDTA blood samples by extraction of white blood cells with a Biomek FX system (Beckman Coulter) using the Nucleo-Mag® Blood 3 ml Kit (Machery-Nagel, #REF 744502.1) as per manufacturer’s instructions. All DNA samples showed high purity as determined by optical density measurements ($A_{260}/A_{280} > 1.9$ and $A_{260}/A_{230} > 2.0$). The diagnosis of MND was made exclusively by experienced neurologists in outpatient clinics specialized in motor neuron disease or neuromuscular diseases. The corresponding diagnoses were as follows: ALS according to Gold Coast Criteria ($n = 76$; 71.0%), PLS according to Turner et al. ($n = 16$; 15.0%), progressive muscular atrophy, PMA ($n = 7$; 6.5%), MND ($n = 8$; 7.5%) (see Suppl. Figure 1) [20, 21]. PMA was diagnosed after thorough exclusion of alternative diagnosis (i.e., polyneuropathy, SBMA, adult-onset SMA, dHMN, and myopathy).

Repeat analysis of *RFC1*, *C9orf72*, and *AR* by Oxford nanopore technology long-read sequencing

All patients were analyzed by an Oxford nanopore technology long-read sequencing-based repeat analysis to identify pathogenic repeat expansions in *C9orf72*, *RFC1* and *AR* for Kennedy’s disease. Library preparation and flow cell loading were performed according to the Oxford nanopore technology (ONT) Cas9-targeted sequencing protocol using ONT’s SQK-CS9109 kit and 5 μ g of input gDNA. CRISPR RNAs (crRNAs) to enrich *C9orf72*, *RFC1*, and *AR* were designed using CHOPCHOP 8 (Supplementary Table 1). Sequencing was performed with ONT FLO-MIN106D R9 flow cells on the GridION X5 sequencer.

Data were analyzed as previously described [4]. Base calling from electrical data was performed using Guppy (v5.0.16) [22, 23]. The generated FASTQ files were aligned to the human reference genome (GRCh38/hg38) using Minimap2 (v2.17) to identify the reads spanning the targets of interest [24]. For quality control of the aligned reads,

we used NanoPlot (v1.29.1) [25]. The bioinformatics tool STRique (v0.2.1) was used to determine the number of repeat units (RU) for all reads assigned to *C9orf72* [26]. Repeat size distributions obtained by STRique were visualized as violin plots and used to determine the repeat size for each allele by computing local maxima using FindPeaks (v2.1.1) and visual assessment of the plots [27]. For extended repeat expansion (> 100 RU) showing multiple local maxima, a size range was calculated starting with the lowest and ending with the highest significant peak as determined by FindPeaks. Only fragments spanning the entire repeat were considered for the repeat length quantification. The *RFC1* locus was analyzed manually by visual inspection of all reads mapped to that region in the IGV. Repeat lengths distribution as well as average repeat sizes were extracted for each allele separately. Patients being heterozygous for a pathogenic repeat expansion in *RFC1* were analyzed for a pathogenic variant in *trans* by NGS sequencing as rare cause of *RFC1* spectrum disorder.

High-throughput sequencing and bioinformatics pipeline

To identify pathogenic single nucleotide variants (SNVs) or copy number variations (CNVs) in known, altogether 380 MND-related genes, gene targeted enrichment was performed with the SureSelectXT gene panel custom kit (Agilent Biosciences) or the Twist Human Comprehensive Exome Kit (Twist Biosciences). Massively parallel sequencing was carried out on an Illumina NextSeq 500 or a Novaseq 6000 system (Illumina, San Diego, CA) as 150 bp paired-end runs using v2.0 SBS chemistry. Secondary and tertiary analysis was carried out using varvis® 1.22.0. Pipeline versions fc9-063-00b and 8a2-0c8-080 were used for SNV and CNV analysis, respectively. The 380 MND-related genes analyzed are summarized in Table 2 of the Supplement. In addition, *RFC1* was analyzed in all patients with a heterozygous pathogenic repeat expansion in *RFC1*. Homozygous deletion in *SMN1* was excluded by a masking pipeline as previously described [28].

Only SNVs and small insert and deletions (INDELs) in the coding and flanking intronic regions (± 50 bp) were evaluated. Variants were classified according to the ACMG (American College of Medical Genetics and genomics) guidelines [29]. Likely pathogenic and pathogenic variants are summarized as P/LP variants.

Data availability

Anonymized data from this study are available from the corresponding author on reasonable request.

Results

Identification of known MND-related genetic alterations

To identify known genetic alterations associated with MND, all 107 patients underwent parallel repeat analysis by ONT long-read sequencing of the ALS-related *C9orf72* repeat array and the *AR* locus associated with spinal and bulbar muscular atrophy (SBMA/Kennedy's disease) as well as NGS sequencing of MND-related genes (Fig. 1, Supplementary Table 3) [20]. Overall, in 14 individuals (13%), a related genetic alteration could be identified. In 8/107 patients (7.5%), a pathogenic repeat expansion in *C9orf72* was detected [30]. None of the patients carried a pathogenic repeat expansion in *AR* associated with SBMA. Heterozygous P/LP variants in genes associated with ALS were identified in six patients: Three patients carried a variant in *SOD1*, with one patient additionally carrying a variant in *FIG4*. One patient each had a variant in *TARDBP*, in *VCP* or in *NEK1*, respectively. In addition, five patients had a variant of uncertain significance in ALS-related genes (*FUS*, *TBKI*, *SETX*, and *SOD1*; Supplementary Table 3). Neither P/LP nor variants of unclear significance (VUS) were found in any other of the altogether analyzed 380 genes associated with MND besides classical ALS with affection of only either the upper or lower motor neurons. Thus, in the majority of patients, no pathogenic or likely pathogenic variant or repeat expansion in a known ALS-related gene or genes associated with MND could be identified.

In-depth analysis of the RFC1 repeat array in MND patients

RFC1 repeat configurations (repeat motifs and repeat lengths) were determined for all patients (Fig. 1b, c, Supplementary Table 3) in parallel with *C9orf72* repeat analysis by ONT long-read sequencing. In three of the 107 MND patients (i.e., 2.8%), a biallelic pathogenic (AAGGG)_{exp} repeat expansion in *RFC1* was detected, one of which was previously reported by us [15]. No other genetic variants in any of the MND-related genes were detected in the panel analyses in these individuals. In all three individuals, the repeat length exceeded the currently established pathogenic cut-off of 250 repeat units [3]. Eight patients carried a pathogenic repeat configuration on one allele, but a non-pathogenic one on the other allele. In all these patients, pathogenic SNVs in *RFC1* in *trans* were excluded by NGS sequencing of *RFC1*.

We examined the frequency of the different repeat motifs (Fig. 1b) in *RFC1* and their repeat sizes (Fig. 1c)

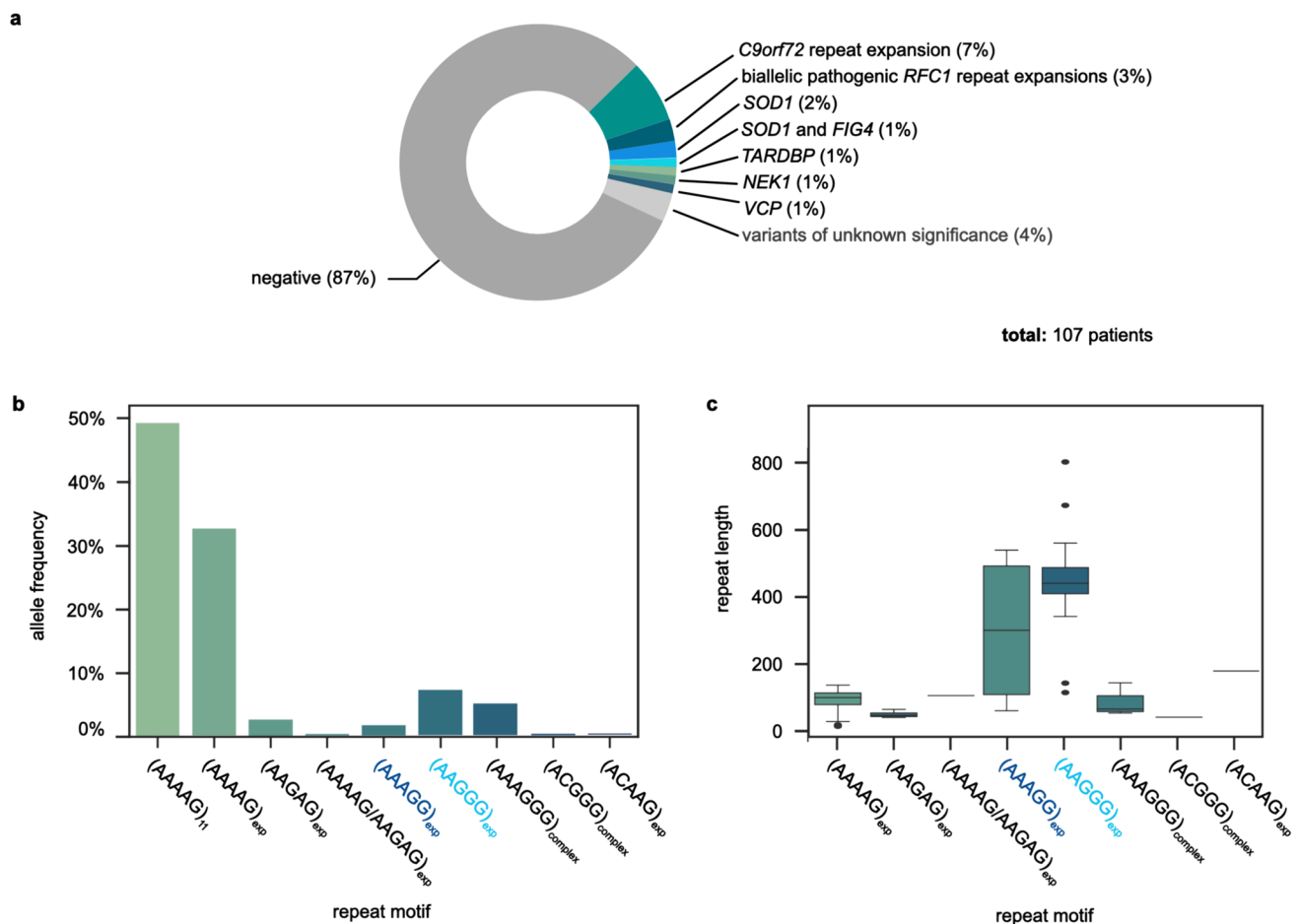


Fig. 1 Results of the genetic analysis of 107 patients with MND. (a) Genetic causes of MND identified by long-read sequencing-based repeat analysis and NGS sequencing of MND-related genes. (b) Repeat motifs of the *RFC1* repeat array identified in all patients ana-

lyzed and their frequency. Pathogenic motif in light blue. Motif with size-dependent pathogenicity in dark blue. (c) Repeat length distribution of the different repeat motifs of the *RFC1* locus determined in the cohort studied

within the 214 alleles of our 107 patients. The unexpanded *RFC1* wild-type motif $(AAAAG)_{11}$ was observed in 50% of the alleles. The expanded non-pathogenic wild-type motif $(AAAAG)_{exp}$ was found in 32% of the alleles, ranging from 16 to 138 RU in size. The likely non-pathogenic motif $(AAGAG)_{exp}$ was found in 2% of the alleles with a small variation in size from 42 to 66 RU. In addition, one polymorphic allele was found in which $(AAGAG)_{exp}$ was surrounded by a stretch of $(AAAAG)_{exp}$ at both sides. The motif with size-dependent pathogenicity $(AAAGG)_{exp}$ was detected in 1% of the alleles with very different repeat lengths (126 RU, 476 RU and 539 RU). Considering the currently estimated threshold for pathogenicity of 600 RU for this motif, we classified all $(AAAGG)_{exp}$ alleles detected as non-pathogenic [3]. The known pathogenic motif $(AAGGG)_{exp}$ was detected in 7% of the alleles, ranging in size from 116 to 800 RU. Two of the alleles showed

a size below the currently established threshold of pathogenicity of 250 RU [3].

In addition, complex polymorphic repeat configurations were identified in approximately 6% of the alleles. Intermittent sequences of $(AAAGGG)_{exp}$ interrupted by $(AAAGG)_{exp}$ and/or $(AAGGG)_{exp}$ were detected in heterozygosity in 11 patients, all without a second pathogenic expansion on the other allele. One patient showed a complex $(ACGGG)_{exp}$ repeat expansion interrupted by stretches of $(AAAGG)_{exp}$ and $(AAGGG)_{exp}$ in one allele, which has previously been reported as of unknown pathogenicity [3, 7]. However, this individual also carried a pathogenic repeat expansion in *C9orf72*. The two types of complex repeat expansions were annotated as $(AAAGGG)_{complex}$ and $(ACGGG)_{complex}$. In addition, we could identify the yet undescribed repeat configuration $(ACAAG)_{exp}$ with a size of 180 RU heterozygous in one patient carrying the wild-type motif $(AAAAG)_{11}$ on the other allele (Supplementary Table 3, Supplementary Fig. 2).

Clinical presentation and diagnostic findings of MND patients with biallelic pathogenic repeat expansions in RFC1

The three male MND patients with a biallelic pathogenic repeat expansion in *RFC1* had a mean age of 59 years and a mean age of onset of 52 years. Two patients (patient #87 and #99) were diagnosed clinically with ALS fulfilling existing Gold Coast Criteria, while the other patient (patient #98) was diagnosed with definite PLS according to existing diagnostic criteria for PLS [20, 21]. The family history for MND was negative in all of them. In the two individuals with ALS, simultaneous and clear clinical signs of upper and lower motor neuron affection were detected including brisk and pathological tendon reflexes, as well as muscle weakness and atrophy with muscle fasciculations (Table 1). Both patients had a spinal onset of symptoms with one of them developing bulbar palsy with dysphagia, as well as dysarthria during disease progression. The third patient (patient #98) primarily presented with upper motor

neuron affection manifesting with spasticity and pseudobulbar palsy fulfilling the diagnostic criteria for definite PLS [21]. This patient had no clinical signs of lower motor neuron affection (i.e., muscle atrophy and/or polytopic muscle fasciculations). By needle electromyography, there were only unspecific findings with single fasciculation potentials and mild to moderate chronic denervation in the anterior tibial muscle but no positive sharp waves and/or fibrillation potentials and no pathological findings in the examined muscles of the other body regions (bulbar–cervical–thoracic). Relevant differential diagnoses were excluded by magnetic resonance imaging of the brain and spinal cord (exclusion of lesions along the pyramidal tracts, brain stem and cervical/thoracic spinal cord), additional laboratory tests, exclusion of genetic alterations in MND-related genes, and particularly thorough electroneurography for exclusion of demyelinating immune neuropathy. Neurofilament light chain levels were analyzed in 2/3 patients either in serum or in cerebrospinal fluid, respectively, and found to be significantly elevated (see Table 1). Two patients showed autonomic dysfunction. After

Table 1 Symptoms of three patients with *RFC1* spectrum disorder and a clinical main phenotype of MND

| Patient | 87 | 98 | 99 |
|-------------------------------------|--|--|---|
| Gender | Male | Male | Male |
| Age | 66 | 67 | 43 |
| Age of onset | 63 | 54 | 41 |
| Affection of the upper motor neuron | Brisk muscle reflexes (cervical) | Brisk muscle reflexes (cervical, lumbosacral) Paraspasticity (lumbosacral) Pseudobulbar palsy (dysphagia and dysarthria) | Brisk muscle reflexes (cervical, lumbosacral) |
| Affection of the lower motor neuron | Muscle weakness and atrophy (cervical, thoracic, lumbosacral) Fasciculations (all 4 levels) | None | Muscle weakness and atrophy (cervical, lumbosacral) Fasciculations (lumbosacral) Bulbar palsy (dysphagia, dysarthria) |
| EMG | Acute denervation in all 4 levels | Chronic neurogenic changes Fasciculations in tibialis anterior muscle | No data available |
| Gold Coast Criteria (2020) | Fulfilled | not fulfilled | Fulfilled |
| Turner criteria | Not fulfilled | Fulfilled (definite) | Not fulfilled |
| MRI (brain, spinal cord) | No relevant findings | Cerebral atrophy, focal white matter lesions, thoracic spinal cord hypotrophy without lesions | No data available |
| <i>Neurofilaments:</i> | | | |
| NfL, serum | 82 pg/mL | No data available | No data available |
| NfL, liquor | No data available | 4260 pg/mL | No data available |
| pNfH, liquor | No data available | 863 pg/mL | No data available |
| Chronic cough | – | + | + |
| sensory PNP | +(Subclinically; axonal) | +(Demyelinating) | +(Axonal) |
| Cerebellar ataxia | – | + | – |
| Bilateral vestibulopathy | +(Presbyvestibulopathy) | + | – |
| Autonomic dysfunction | – | + | + |

identification of biallelic pathogenic repeat expansions in *RFC1*, reverse phenotyping revealed additional symptoms of *RFC1* spectrum disorder in all patients. As such chronic cough (2/3 patients), sensory neuropathy (3/3 patients), cerebellar ataxia (1/3), as well as bilateral vestibulopathy (2/3) could be identified. While the progression of the disease in the patient previously described (patient #87) resembled that of a typical ALS course of disease, the two other patients showed a rather slow progression of motor symptoms: in one patient, they manifested more than a decade ago (patient #98) with an almost stable to slowly progressive course of disease with a late onset of chronic neurogenic changes as well as slight fasciculations and a similarly slow course of disease in the other patient for about 2 years (patient #99).

Discussion

Prompted by the discovery of pathogenic biallelic repeat expansions in *RFC1* in an ALS patient with additional sensory neuropathy and bilateral vestibular dysfunction [15], we investigated whether biallelic pathogenic repeat expansions in *RFC1* could be another relevant genetic cause for motor neuron disease such as ALS, PLS or PMA. Therefore, we systematically characterized the *RFC1* repeat array in a cohort of 107 patients with clinical diagnosis of MND (Fig. 1b, c, Supplementary Table 3). For the first time in such a cohort, repeat motifs and accurate repeat lengths of *RFC1* were determined in parallel and together with repeat expansions in *C9orf72* and *AR* by an ONT long-read sequencing method. Pathogenic biallelic repeat expansions in *RFC1* in absence of another causative genetic alteration were identified in 3% of our patients, including the patient originally identified and reported as a case report [15]. Assuming a heterozygous carrier frequency of the pathogenic ‘AAGGG’ repeat motif of up to 4% in the European population resulting in an estimated prevalence of *RFC1* spectrum disorder of up to 1/2500, individuals with a biallelic pathogenic repeat expansion were enriched by the factor 70 in this cohort, making a coincidence highly unlikely [1, 2, 31].

Interestingly, a recent study by Abramzon et al. investigated whether biallelic pathogenic repeat expansions in *RFC1* can be detected in ALS patients. When screening a cohort of 1069 patients with a clinical diagnosis of sporadic ALS, in contrast to our study, no patient with biallelic pathogenic repeat expansions in *RFC1* was found [6]. A possible explanation for these discrepant results could be the composition of the cohorts and different methods used for analyzing the *RFC1* repeat array: The study of Abramzon et al. only included ALS patients from a study registry fulfilling the El Escorial criteria, whereas in this study patients with a clinical diagnosis of MND (primarily ALS) from a real-world clinical outpatient cohort were analyzed.

Another possible explanation for the identification of two patients with a clinical diagnosis of ALS in this study could be the known geographic differences in the frequencies of known ALS-associated genetic variants, even across populations with European ancestry [32]. In addition, the applied diagnostic approaches differed significantly. Abramzon et al. used an iterative PCR-based workflow without accurately determining the repeat configuration of both alleles employing ONT long-read sequencing. This might have missed patients with an *RFC1* spectrum disorder especially given the limited knowledge about the pathogenicity of individual motifs or a potential pathogenic SNV in individuals heterozygous for a pathogenic repeat expansion at that time [3, 33].

Initially, biallelic pathogenic repeat expansions in *RFC1* were identified in patients with CANVAS, a characteristic clinical triad of cerebellar ataxia, sensory neuropathy and vestibular areflexia [1, 2]. Meanwhile, incomplete CANVAS phenotypes and a variety of additional, atypical symptoms such as autonomic dysfunction, bradykinesia, parkinsonism or dystonia expanded the phenotype of patients carrying pathogenic repeat expansions, which led to the terminology of *RFC1* spectrum disorder [9, 34, 35]. The severity of the individual clinical presentation varies widely, and in some patients, might even fully resemble another disease phenotype, such as multiple system atrophy of the cerebellar type or parkinsonism [12, 36].

Recently, motor neuron affection has been reported as frequent additional symptom of *RFC1* spectrum disorder [13, 14]. As such, one study identified predominant affection of either the upper motor neuron in 29% or the lower motor neuron in 18% of *RFC1* patients and simultaneous affection of both, the upper and the lower motor neuron, in 16% of *RFC1* patients [13]. Furthermore, as a direct morphological correlate for clinical signs of motor neuron affection, post-mortem neuropathological examination of one patient with *RFC1* spectrum disorder showed moderate but significantly increased motor neuron loss, particularly in the anterior horn of the thoracic spinal cord and brainstem motor nuclei such as the hypoglossal nucleus. A potential mechanism of motor neuron affection is axonal swelling between the upper and lower motor neuron, which was observed in one *RFC1* patient thus leading to synaptic dysfunction [13]. This pathomechanism seems to be different to that usually observed in patients with classical ALS, which is characterized by cytoplasmic inclusions of abnormally aggregated and posttranslationally modified TAR DNA binding protein (TDP-43) in neuronal cells. Postmortem neuropathological examinations of two individuals with *RFC1* spectrum disorder and motor neuron affection have shown the absence of such inclusions [35]. Thus, it seems to be the case that different pathomechanisms besides classical intraneuronal TDP-43 pathology could lead to motor neuron loss manifesting with similar

clinical phenotypes in *RFC1*-pathology, but obviously with a slower progression rate as compared to classical ALS.

Besides being the origin of the distinct phenotype of CANVAS, biallelic pathogenic repeat expansions in *RFC1* might additionally induce neurodegeneration of other functional systems via yet unknown pathways. This hypothesis is in line with the recent discussion of clinical and pathological pleiotropy of neurodegenerative diseases, where similar genetic alterations can cause different phenotypes and vice versa [37]. Independent of biallelic ‘AAGGG’ repeat expansions that cause a neurodegenerative phenotype such as CANVAS with high penetrance, other *RFC1* repeat configurations could be potential genetic risk factors for a neurodegeneration such as ALS as it is known for certain intermediate and pathogenic repeat expansions such as in *ATXN1*, *ATXN2* and *HTT* [37–39]. Furthermore, biallelic pathogenic *RFC1* expansion might explain previous reports and studies with growing evidence on sensory and cerebellar involvement in ALS [40–42].

We statistically evaluated the configuration of the *RFC1* repeat array in MND patients to study its variability and a potential enrichment of distinct patterns in this cohort. However, the identification or exclusion of specific repeat configurations as risk factor for MND cannot be performed in this study. This is due to the rather small impact of polygenic risk factors on the overall disease risk, which requires larger cohorts of MND patients as well as a cohort of healthy individuals characterized by long-read sequencing to be detected.

Instead, our data highlight the intra- and interallelic heterogeneity of the *RFC1* repeat array with 6% of the alleles showing complex repeat patterns, that cannot be detected by standard PCR-based repeat analysis. Using ONT long-read sequencing for the analysis, we were able to identify what we believe to be a novel repeat motif (ACAAG)_{exp} with a size of 180 RU. Recently, it has been hypothesized that the combination of GC-content and repeat size of a certain repeat configuration determines its pathogenicity [3]. As the GC-content in ‘ACAAG’ is lower compared to known pathogenic repeat motifs and the detected repeat length is rather short, we postulate a non-pathogenicity of the found repeat configuration. In summary, we present repeat expansions in *RFC1* in about 3% of individuals with a clinical main phenotype of MND (i.e., two cases with ALS according to existing Gold Coast Criteria and one patient with PLS, respectively upper motor neuron dominant ALS), emphasizing that a pure or predominant motor neuron disease might be another extreme phenotype of *RFC1* spectrum disorder (see Suppl. Figure 3) and thus should be considered in the genetic diagnostic workup in patients diagnosed with MND, especially when other additional symptoms like sensory and/or autonomic neuropathy, vestibular failure, cerebellar deficits and/or a chronic cough occur. For an in-depth analysis

of the heterogeneous *RFC1* locus including an accurate determination of its repeat size and motif, it is crucial to employ long-read sequencing, offering an unbiased approach to adequately capture and understand the complex genetics of *RFC1* spectrum disorder. However, further studies are needed to validate our findings and to definitely assess the role of pathogenic motifs and repeats in *RFC1* for the upper and lower motor neurons as well as to analyze a possible pathogenic role of novel repeat motifs with regard to a polygenic risk.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00415-024-12519-6>.

Acknowledgements A.S. is grateful for a scholarship from the Medical Faculty of the University of Munich (LMU).

Author contributions A.S., H.E., A.A., P.R. and F.S. conceptualized and designed the study. M.T., R.G., P.R. and F.S. recruited the patients of this study. R.G., P.R. and F.S. planned and performed clinical diagnostics and interpreted the results. A.S., H.E., M.T., V.S. and A.A. planned and performed the genetic testing and interpreted results. A.S., H.E., I.C., A.A., P.R. and F.S. wrote the manuscript. All the authors revised the manuscript, provided feedback, and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Competing interests FS reports honoraria for advisory boards for Amylyx, Alnylam and Alexion. All the other authors report no competing interest.

Ethical standards Informed consent was obtained from all participants. All genetic analyses and investigations were performed in accordance with the guidelines of the Declaration of Helsinki and approved by local institutions (Bayerische Landesärztekammer, vote no. 2019-210).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Cortese A, Simone R, Sullivan R et al (2019) Biallelic expansion of an intronic repeat in *RFC1* is a common cause of late-onset ataxia. *Nat Genet* 51:649–658. <https://doi.org/10.1038/s41588-019-0372-4>
2. Rafehi H, Szmulewicz DJ, Bennett MF et al (2019) Bioinformatics-based identification of expanded repeats: a non-reference

- intronic pentamer expansion in RFC1 causes CANVAS. *Am J Hum Genet* 105:151–165. <https://doi.org/10.1016/j.ajhg.2019.05.016>
3. Dominik N, Magri S, Currò R et al (2023) Normal and pathogenic variation of RFC1 repeat expansions: implications for clinical diagnosis. *Brain* 146:5060–5069. <https://doi.org/10.1093/brain/awad240>
 4. Erdmann H, Schöberl F, Giurgiu M et al (2023) Parallel in-depth analysis of repeat expansions in ataxia patients by long-read sequencing. *Brain* 146:1831–1843. <https://doi.org/10.1093/brain/awac377>
 5. Scriba CK, Beecroft SJ, Clayton JS et al (2020) A novel RFC1 repeat motif (ACAGG) in two Asia-Pacific CANVAS families. *Brain J Neurol* 143:2904–2910. <https://doi.org/10.1093/brain/awaa263>
 6. Abramzon Y, Dewan R, Cortese A et al (2021) Investigating RFC1 expansions in sporadic amyotrophic lateral sclerosis. *J Neurol Sci* 430:118061. <https://doi.org/10.1016/j.jns.2021.118061>
 7. Igor S, Chintalaphani SR, Hasindu G et al (2022) Comprehensive genetic diagnosis of tandem repeat expansion disorders with programmable targeted nanopore sequencing. *Sci Adv* 8:5386. <https://doi.org/10.1126/sciadv.abm5386>
 8. Cortese A, Tozza S, Yau WY et al (2020) Cerebellar ataxia, neuropathy, vestibular areflexia syndrome due to RFC1 repeat expansion. *Brain* 143:480–490. <https://doi.org/10.1093/brain/awz418>
 9. Traschütz A, Cortese A, Reich S et al (2021) Natural history, phenotypic spectrum, and discriminative features of multisystemic RFC1 disease. *Neurology* 96:e1369–e1382. <https://doi.org/10.1212/WNL.00000000000011528>
 10. Kytövuori L, Sipilä J, Doi H et al (2022) Biallelic expansion in RFC1 as a rare cause of Parkinson's disease. *NPJ Park Dis* 8:6. <https://doi.org/10.1038/s41531-021-00275-7>
 11. da Silva SG, Martinez ARM, da Graça FF et al (2020) Dopamine responsive parkinsonism in a patient with homozygous RFC1 expansions. *Mov Disord* 35:1889–1890. <https://doi.org/10.1002/mds.28286>
 12. Wan L, Chen Z, Wan N et al (2020) Biallelic intronic AAGGG expansion of RFC1 is related to multiple system atrophy. *Ann Neurol* 88:1132–1143. <https://doi.org/10.1002/ana.25902>
 13. Huin V, Coarelli G, Guemy C et al (2022) Motor neuron pathology in CANVAS due to RFC1 expansions. *Brain* 145:2121–2132. <https://doi.org/10.1093/brain/awab449>
 14. Reyes-Leiva D, Aldecoa I, Gelpi E, Rojas-García R (2022) Motor neuron involvement expands the neuropathological phenotype of late-onset ataxia in RFC1 mutation (CANVAS). *Brain Pathol* 32:e13051. <https://doi.org/10.1111/bpa.13051>
 15. Schoeberl F, Abicht A, Kuepper C et al (2022) Sensory neuropathy due to RFC1 in a patient with ALS: more than a coincidence? *J Neurol* 269:2774–2777. <https://doi.org/10.1007/s00415-021-10835-9>
 16. Lattante S, Conte A, Zollino M et al (2012) Contribution of major amyotrophic lateral sclerosis genes to the etiology of sporadic disease. *Neurology* 79:66–72. <https://doi.org/10.1212/WNL.0b013e31825dceca>
 17. Ruf WP, Boros M, Freischmidt A et al (2023) Spectrum and frequency of genetic variants in sporadic amyotrophic lateral sclerosis. *Brain Commun* 5:152. <https://doi.org/10.1093/braincomms/fcad152>
 18. de Boer EMJ, de Vries BS, Pennings M et al (2023) Genetic characterization of primary lateral sclerosis. *J Neurol* 270:3970–3980. <https://doi.org/10.1007/s00415-023-11746-7>
 19. Gilpatrick T, Lee I, Graham JE et al (2020) Targeted nanopore sequencing with Cas9-guided adapter ligation. *Nat Biotechnol* 38:433–438. <https://doi.org/10.1038/s41587-020-0407-5>
 20. Shefner JM, Al-Chalabi A, Baker MR et al (2020) A proposal for new diagnostic criteria for ALS. *Clin Neurophysiol* 131:1975–1978. <https://doi.org/10.1016/j.clinph.2020.04.005>
 21. Turner MR, Barohn RJ, Corcia P et al (2020) Primary lateral sclerosis: consensus diagnostic criteria. *J Neurol Neurosurg Psychiatry* 91:373–377. <https://doi.org/10.1136/jnnp-2019-322541>
 22. Wick RR, Judd LM, Holt KE (2019) Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol* 20:129. <https://doi.org/10.1186/s13059-019-1727-y>
 23. Oxford Nanopore Technologies. Guppy protocol: modified base calling. https://community.nanoporetech.com/protocols/Guppy-protocol/v/gpb_2003_v1_revz_14dec2018/modified-base-calling. Accessed 02 Feb 2022.
 24. Li H (2018) Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100
 25. De Coster W, D'Hert S, Schultz DT et al (2018) NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
 26. Giesselmann P, Brändl B, Raimondeau E et al (2019) Analysis of short tandem repeat expansions and their methylation state with nanopore sequencing. *Nat Biotechnol* 37:1478–1481. <https://doi.org/10.1038/s41587-019-0293-x>
 27. (2022) GitHub. In: GitHub. <https://github.com/erdogant/findpeaks>. Accessed 2 Feb 2022
 28. Kleinle N, Scholz V, Benet-Pagés A et al (2023) Closing the gap—detection of 5q-spinal muscular atrophy by short-read next-generation sequencing and unexpected results in a diagnostic patient cohort. *J Neuromuscul Dis* 10:835–846. <https://doi.org/10.3233/JND-221668>
 29. Richards S, Aziz N, Bale S et al (2015) 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* 17:405–423. <https://doi.org/10.1038/gim.2015.30>
 30. Renton AE, Majounie E, Waite A et al (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72:257–268. <https://doi.org/10.1016/j.neuron.2011.09.010>
 31. Cortese A, Reilly, Mary M, Houlden, Henry (2021) RFC1 CANVAS / Spectrum Disorder. In: GeneReviews®. <https://www.ncbi.nlm.nih.gov/books/NBK564656/>. Accessed 28 Apr 2021
 32. Al-Chalabi A, Hardiman O (2013) The epidemiology of ALS: a conspiracy of genes, environment and time. *Nat Rev Neurol* 9:617–628. <https://doi.org/10.1038/nrneurol.2013.203>
 33. Ronco R, Perini C, Currò R et al (2022) Truncating variants in RFC1 in cerebellar ataxia, neuropathy, and vestibular areflexia syndrome. *Neurology*. <https://doi.org/10.1212/WNL.0000000000201486>
 34. Malaquias MJ, Braz L, Santos Silva C et al (2023) Multisystemic RFC1-related disorder: expanding the phenotype beyond cerebellar ataxia, neuropathy, and vestibular areflexia syndrome. *Neurol Clin Pract* 13:e200190. <https://doi.org/10.1212/CJP.0000000000200190>
 35. Ylikotila P, Sipilä J, Alapirtti T et al (2023) Association of biallelic RFC1 expansion with early-onset Parkinson's disease. *Eur J Neurol* 30:1256–1261. <https://doi.org/10.1111/ene.15717>
 36. Davies K, Szmulewicz DJ, Corben LA et al (2022) RFC1-related disease: molecular and clinical insights. *Neurol Genet* 8:e200016. <https://doi.org/10.1212/NXG.0000000000200016>
 37. Henden L, Fearnley LG, Grima N et al (2023) Short tandem repeat expansions in sporadic amyotrophic lateral sclerosis and frontotemporal dementia. *Sci Adv* 9:2044. <https://doi.org/10.1126/sciadv.ade2044>
 38. Lattante S, Pomponi MG, Conte A et al (2018) ATXN1 intermediate-length polyglutamine expansions are associated with

- amyotrophic lateral sclerosis. *Neurobiol Aging* 64:157.e1-157.e5. <https://doi.org/10.1016/j.neurobiolaging.2017.11.011>
39. Naruse H, Matsukawa T, Ishiura H et al (2019) Association of ATXN2 intermediate-length CAG repeats with amyotrophic lateral sclerosis correlates with the distributions of normal CAG repeat alleles among individual ethnic populations. *Neurogenetics* 20:65–71. <https://doi.org/10.1007/s10048-019-00570-9>
 40. Kleinerova J, Tahedl M, Tan EL et al (2024) Supra- and infratentorial degeneration patterns in primary lateral sclerosis: a multimodal longitudinal neuroradiology study. *J Neurol* 271:3239–3255. <https://doi.org/10.1007/s00415-024-12261-z>
 41. Chipika RH, Mulkerrin G, Pradat P-F et al (2022) Cerebellar pathology in motor neuron disease: neuroplasticity and neurodegeneration. *Neural Regen Res* 17:2335–2341. <https://doi.org/10.4103/1673-5374.336139>
 42. Pugdahl K, Fuglsang-Frederiksen A, de Carvalho M et al (2007) Generalised sensory system abnormalities in amyotrophic lateral sclerosis: a European multicentre study. *J Neurol Neurosurg Psychiatry* 78:746–749. <https://doi.org/10.1136/jnnp.2006.098533>