Repeat-associated ataxias in a German patient cohort analysed by targeted parallel long-read sequencing

Hannes Erdmann,[1,2] Annalisa Schaub,[1,2] Morghan C. Lucas,[1,2,3] Veronika Scholz,[1] Anna Benet-Pages,[1,4] Kerstin Becker,[1] Christine Dineiger,[1] Veronika Mayer,[1] Inga van Buren,[1] Eva Breithausen,[1] Karl Akbari,[5] Isabell Cordts,[6,7] Mayra Sauer,[1] Christine Schneider,[8] Rosanna Krakowsky,[9] Franziska Schnabel,[9] Konstanze Dunker,[10] Lena Fabritius,[10] Johannes Gerb,[10,11] Denis Grabova,[10] Ken Möhwald,[10,11] Marius Näher,[10] Karoline Steinmetz,[10] Franziska Thiessen,[10] Alexander Jäck,[11,12] Christiane Schneider-Gold,[13] Simone Zittel,[14] Christina Petersen,[15] Isolde Schreyer,[15] Larissa Mämecke,[16] Sibylle Wilfling,[17] Gilbert Wunderlich,[18] David Brenner,[19,20,21] Yorck Hellenbroich,[22] Kirsten Muhle,[23] Tessa Huchtemann,[23] Inga Claus,[23] Thomas Klopstock,[2,12,24] Michael Strupp,[11] Johannes Levin,[11,12,24] Günter Höglinger,[11,12,24] Doreen Huppert,[10] Sandra Becker-Bense,[10] Filipp Filippopulos,[10,11] Fabian Kilpert,[25] Elsa Leitão,[25] Sabine Kaya,[25] Christel Depienne,[25] Florian Schöberl,[11] Teresa Neuhann,[1] Elke Holinski-Feder,[1,3] Andreas Zwergal[10,11] and Angela Abicht[1,2]

Abstract

Hereditary adult-onset ataxias are a heterogeneous group of phenotypically overlapping conditions, often caused by pathogenic expansions of short tandem repeats. Currently, 18 repeat disorders with a core phenotype of adult-onset ataxia are known. Diagnosis typically relies on sequential PCR-based methods, which are labour-intensive and lack precision. Long-read sequencing (LRS) has the potential to overcome these limitations and is currently implemented and validated in clinical genetics. Using clinical nanopore Cas9-targeted sequencing (Clin-CATS) for parallel in-depth repeat analysis, we evaluated a diagnostic cohort of 513 adult-onset ataxia patients, determining frequencies of all known repeat-associated ataxias except Spinocerebellar ataxia 4 (SCA4), as well as the carrier frequencies for autosomal-recessive disorders, *RFC1* spectrum disorder and Friedreich's ataxia (FRDA). Additionally, phenotypes of

patients with established genetic diagnoses were characterized, especially those of patients living with *RFC1* spectrum disorder and SCA27B.

Repeat-associated ataxias were confirmed in 33.3% of cases, including rare ataxias, such as SCA10, SCA36 and SCA37, alongside as the most prevalent conditions SCA27B and *RFC1* spectrum disorder. Potentially pathogenic expansions in *FGF14* were identified in an additional 4.7% of patients. Testing of another 347 patients for *ZFHX3* expansions linked to SCA4 did not identify any cases. Dual diagnoses were frequent, occuring in 6.4% of patients with repeat-associated ataxia. We confirmed a high *RFC1* spectrum disorder carrier frequency (7.2%) and reclassified certain FXN expansions as likely non-pathogenic, resulting in a lower than estimated carrier frequency for FRDA of 0.8%. We also identified novel repeat configurations in several loci and illustrated the high heterogeneity of repeat expansions in *RFC1*, highlighting it as a potential source of false results when using PCR-based methods. This study underscores the diagnostic advantages of LRS for comprehensive repeat analysis and recommends its adoption as a standard in clinical genetics, replacing Southern blot and PCR-based approaches. Furthermore, based on our findings in a large patient cohort a re-evaluation of existing phenotype-genotype correlations is recommended as well as evaluating additional parameters besides repeat length to improve diagnostic precision of repeat analysis.

Author affiliations:

1 Medical Genetics Center (MGZ) Munich, 80335 Munich, Germany

2 Friedrich Baur Institute at the Department of Neurology, LMU University Hospital, Ludwig-Maximilians-Universität München, 80336 Munich, Germany

3 Department of Medicine IV, LMU University Hospital, Ludwig-Maximilians-Universität München, 80336 Munich, Germany.

4 Institute of Neurogenomics, Helmholtz Center Munich, 85764 Neuherberg, Germany

5 School of International Business, Hochschule Bremen, 28199 Bremen, Germany

6 Department of Neurology, Department of Neurology, Klinikum Rechts der Isar, Technical University of Munich, 81675 Munich, Germany

7 Department of Neuroscience, Mayo Clinic, Jacksonville, FL 32224, USA

8 Department of Neurology, University Medical Center Augsburg, 86156 Augsburg, Germany.

9 Institute of Human Genetics, University of Leipzig Medical Center, 04103 Leipzig, Germany

10 German Center for Vertigo and Balance Disorders, LMU University Hospital, Ludwig-Maximilians-Universität München, 81377 Munich, Germany

11 Department of Neurology, LMU University Hospital, Ludwig-Maximilians-Universität München, 81377 Munich, Germany

12 German Center for Neurodegenerative Diseases (DZNE) Munich, Ludwig-Maximilians-Universität München, 81377 Munich, Germany

13 Department of Neurology, St Josef Hospital, Ruhr-University of Bochum, 44791 Bochum, Germany

14 Department of Neurology, University Medical Center Hamburg-Eppendorf, 20246 Hamburg, Germany

15 Praxis für Humangenetik, Zentrum für ambulante Medizin, Institute of Human Genetics, University of Jena, 07743 Jena, Germany

16 Mitteldeutscher Praxisverbund Humangenetik, 06110 Halle/Saale, Germany

17 Center for Human Genetics Regensburg, 93047 Regensburg, Germany

18 Department of Neurology and Center for Rare Diseases, Faculty of Medicine and University Hospital, University of Cologne, 50937 Cologne, Germany

19 Department of Neurology, Ulm University, 89081 Ulm, Germany

20 German Center for Neurodegenerative Diseases (DZNE) Ulm, 89081 Ulm, Germany

21 Center for Rare Diseases (ZSE) Ulm, Ulm University Hospital Center for Rare Diseases, Ulm, 89081 Ulm, Germany

22 Institute of Human Genetics, University Hospital Schleswig-Holstein, University of Luebeck and Kiel, 23538 Luebeck, Germany

23 Department of Neurology, University Hospital Münster, 48149 Münster, Germany

24 Munich Center for Systems Neurology (SyNergy), 81377 Munich, Germany

25 Institute of Human Genetics, University Hospital Essen, University of Duisburg-Essen, 45122, Essen, Germany

Correspondence to: Prof. Dr. med. Angela Abicht

Medical Genetics Center

Bayerstr. 3–5

80335 Munich

Germany

E-mail: Angela.abicht@mgz-muenchen.de

Running title: Repeat-associated ataxias in Germany

Keywords: repeat expansion disorders; long-read sequencing; hereditary ataxia; repeat analysis; RFC1

Introduction

Adult-onset ataxias are a heterogeneous group of phenotypically overlapping conditions characterised by a progressive lack of coordination leading to stance and gait imbalance, ocular motor disturbance, dysarthria, and dysdiadochokinesia. Additional signs and symptoms, such as neuropathy, basal ganglia dysfunction or cognitive and behavioural changes, may also occur.[1,2] Based on their aetiology, these conditions are classified as acquired, sporadic-degenerative, or hereditary.[3]

The most frequent genetic cause of ataxias is the pathogenic expansion of short tandem repeats. To date, 18 repeat disorders with ataxia as a core phenotype have been identified, encompassing autosomal-dominant, autosomal-recessive, and X-linked inheritance patterns (Supplementary Table 1).[4–6] Due to the deep intronic location of their repeat expansions, two of these disorders,

RFC1 spectrum disorder and spinocerebellar ataxia 27B (SCA27B), have only been described recently, despite their suspected high prevalence, likely accounting for a significant proportion of apparently sporadic degenerative ataxias.[5,7–10] *RFC1* spectrum disorder results from biallelic pathogenic repeat expansions in *RFC1*.[7,8] Initially linked to the clinical triad of cerebellar ataxia, neuropathy and vestibular areflexia syndrome (CANVAS), it is now associated with a broader range of neurological conditions, in particular incomplete presentations of CANVAS and additional symptoms such as chronic cough, autonomic dysfunction and motoneuron involvement.[11–14] SCA27B, an autosomal-dominant ataxia caused by intronic (GAA)-expansions in *FGF14*, is often episodic, triggered by physical exertion, and progresses slowly, complicating its diagnosis, despite the availability of symptomatic treatments, such as 4-aminopyridine.[5,9,15–17]

Hereditary ataxias cannot be reliably diagnosed on clinical findings alone, making molecular genetic testing essential for accurate diagnosis.[18] Standard diagnostic methods, including PCR-based methods and Southern blotting, focus on repeat length as the primary diagnostic parameter. These approaches are limited in their ability to detect complex repeat configurations, where pathogenicity depends on distinct motifs exceeding a specific length threshold, as seen in *RFC1* spectrum disorder, SCA27B, SCA31 and SCA37.[19–21] Although repeat-primed PCR methods can detect the presence or absence of the expansion of a specific motif, they do not provide information on repeat length, nor do they capture all possible repeat motifs or complex repeat configurations, as is often the case for expansions in *RFC1*.[19,22,23]

Long-read sequencing (LRS) has emerged as a transformative method for repeat analysis, offering parallel and detailed analysis of multiple repeat loci while overcoming current diagnostic limitations.[24–26] We recently implemented clinical nanopore Cas9-targeted sequencing (Clin-CATS) for the comprehensive, in-depth repeat analysis of patients with adult-onset ataxia.[25] Initially covering the 10 most common repeat-associated ataxias known in 2020, we have since extended our assay to cover all known repeat-associated adult-onset ataxias. SCA4 was added after the completion of this study, following the recent identification of its molecular origin.[6,27]

In this study, we applied Clin-CATS to a clinical cohort of 513 adult-onset ataxia patients in Germany. We determined the diagnostic yield, assessed the frequency of the different conditions and their genotypes, including configuration (repeat length plus repeat motif) of polymorphic

repeat expansions in *RFC1,* and reassessed the carrier frequencies of the autosomal-recessive disorders, *RFC1* spectrum disorder and Friedreich's ataxia (FRDA). Based on our findings, we discuss challenges in clinically assessing repeat expansions and compare Clin-CATS to other LRS approaches.

Materials and methods

Patient DNA samples and study approval

In total, 513 individuals (>20 years old) with clinical diagnosis of cerebellar ataxia and without established genetic diagnosis were analysed with our updated Clin-CATS workflow. The cohort includes patients without prior genetic diagnostics as well as patients having had a negative result in a previous analysis. Only three families, with a total of 7 individuals, are included in the study; the remaining 506 patients are unrelated. Patients were referred to genetic testing from general neurologists, as well as specialized centers for movement disorders from all over Germany likely resembling the general German population living with cerebellar ataxia. Genetic testing was requested only if there was no suspicion or confirmation of another, alternative aetiology of ataxia (e.g., infectious, autoimmune, paraneoplastic, toxic, or structural) after thorough clinical characterization by board-certified neurologists and diagnostic procedures addressing aforementioned aetiologies. Clinical data were collected retrospectively from medical records. Another 347 patients were analysed with the most recent version of Clin-CATS including guides for *ZHFX3*. All patients were analysed as diagnostic samples within our institute between April 2023 and October 2024. Informed consent was obtained from all patients. The study was approved by local institutions (Bayerische Landesärztekammer, 2019-210). All genetic analyses and investigations were performed in accordance with the guidelines of the Declaration of Helsinki.

Extraction of genomic DNA (gDNA)

Genomic DNA (gDNA) was obtained from total peripheral EDTA blood samples by extraction of white blood cells with a Biomek FX system (Beckman Coulter) using the NucleoMag® Blood 3 ml Kit (Machery-Nagel, 744502.1) as per manufacturer's instructions. All DNA samples

showed high purity as determined by optical density measurements (A260/A280 > 1.9 and A260/A230 > 2.0). The DNA Integrity Number (DIN) was measured on the TapeStation 4200 (Agilent Technologies, G2991BA) as per manufacturer's instructions.

Repeat analysis by Oxford Nanopore Technologies LRS

All patients were analysed using Oxford Nanopore Technology (ONT) LRS-based Clin-CATS as validated and accredited diagnostic method in Germany.[25] Library preparation and flow cell loading were performed according to the ONT Cas9-targeted sequencing protocol using ONT's SQK-CS9109 kit and 5 µg of input gDNA. CRISPR RNAs (crRNAs) for the enrichment of 17 regions of interest were designed using CHOPCHOP 8 (Supplementary Table 2). Sequencing was performed with ONT FLO-MIN106D R9 flow cells on the GridION X5 sequencer. Base calling was performed using Guppy (v5.0.16).[28,29] 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.[30] Minimum criteria to evaluate samples in the diagnostic context was a 50x coverage across all target regions. For *RFC1* and *NOP56*, incomplete reads exceeding the current threshold of pathogenicity of the respective motif were considered for calculating the on-target coverage. For quality control of the aligned reads, NanoPlot (v1.29.1) was used.[31] The bioinformatics tool STRique (v0.2.1) was used to determine the number of repeat units (RU) for all reads assigned to the regions of interest [32]. Repeat length distributions obtained by STRique were visualised 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.[33] 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. In all such cases, the median repeat length was calculated from the mapped reads, provided the repeat length exhibited an approximately Gaussian distribution, indicating sufficient coverage of the respective allele. Repeat configurations were analysed manually by visual inspection of all reads mapped to the respective regions in IGV. Repeat length distribution and median repeat length were extracted for each allele separately. For *ATXN8*/*ATXN8OS,* repeat lengths were determined as the sum of the non-pathogenic proximal $(CTA \bullet TAG)_{exp}$ repeat and the adjunct disease-associated $(CTG \bullet CAG)_{exp}$ repeat as established for PCR-based methods. Patients heterozygous for a

pathogenic repeat expansion in *RFC1* or *FXN* were analysed for a pathogenic variant in *trans* by next-generation sequencing (NGS).

The validity of median repeat lengths for the repeat in *FGF14* was confirmed by comparing our results with those of the previously established sequencing of long-range PCR products using ONT long-read sequencing (LR-PCR amplicon sequencing, Supplementary Table 3, Supplementary Figure 1).[34] Patients were diagnosed as positive for a repeat disorder when they carried a reduced-penetrance or full-penetrance repeat expansion. If more than one repeat expansion was detected, the primary diagnosis was based on the repeat expansion with full penetrance, while additional expansions with reduced penetrance were listed as secondary (and tertiary) diagnoses.

Next-generation sequencing (NGS)

All patients being heterozygous carriers for *RFC1* spectrum disorder and *FRDA* were analyzed by NGS sequencing. Further patients were analyzed, when requested by referring physicians based on non-standardized decision criteria. To identify pathogenic single nucleotide variants (SNVs) or copy number variations (CNVs) in known genes associated with ataxia (altogether 250 genes, Supplementary Table 4), 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. 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/AMP (American College of Medical Genetics and Genomics and the Association for Molecular Pathology) guidelines [35]. In total, 185 patients were analysed, comprising 146 with a negative result (139 patients) or only an intermediate *FGF14* allele (7 patients) and 39 with likely causative repeat expansions.

Results

Diagnosis of 513 ataxia patients by parallel LRS-based repeat analysis

We analysed a cohort of 513 patients with adult-onset ataxia (median age 61 years; range: 22–88, Supplementary Table 5) using our updated version of Clin-CATS.[25] This assay captured all known repeat-associated hereditary ataxias, except SCA4, whose molecular origin was only recently discovered (Figure 1, Supplementary Table 1).[6,27] Genetic analysis revealed pathogenic repeat expansions confirming hereditary ataxia in 171 patients (33.3%) (Figure 2, Supplementary Table 5). An additional 24 patients (4.7%) harboured *FGF14* expansions (180-249 RU), which, while potentially pathogenic with reduced penetrance for SCA27B, remain unconfirmed and are henceforth referred to as 'intermediate *FGF14* alleles'.[16,34] The most common repeat expansions were compatible with *RFC1* spectrum disorder (42 patients, 8.2%) and SCA27B (in total 83 patients (16.2%): 55 (10.7%) with $(GAA)_{>300\ RU}$ expansions, 28 (5.4%) with $(GAA)_{250–300\ RU}$ expansions). Including intermediate *FGF14* alleles, 107 patients (20.9%) had results consistent with SCA27B. Pathogenic repeat expansions compatible with a primary diagnosis of SCA8 were the third most common finding (11 patients, 2.1%), followed by SCA3 (7 patients, 1.4%), SCA6 (6 patients, 1.2%), Fragile-X-associated-tremor/ataxia syndrome (FXTAS, 6 males, 1 female; 1.4%), and other rarer expansions: SCA1 (4 patients, 0.8%), FRDA (3 patients, 0.6%), SCA2 (2 patients, 0.4%), SCA10 (2 patients, 0.4%), DRPLA (1 patient, 0.2%), SCA36 (1 patient, 0.2%) and SCA37 (1 patient, 0.2%). One patient exhibited a *TBP* expansion with reduced penetrance alongside a variant of uncertain significance (VUS) in *STUB1*, indicating SCA17 and/or SCA48, with the intermediate allele in *TBP* acting as a disease modifier.[36] No cases of SCA7, SCA12 or SCA31 were detected in this cohort.

Testing of another 347 patients with adult-onset ataxia for SCA4

After completing the study of the 513 patients, another 347 patients were tested for SCA4 for an initial estimate of its frequency in patients with adult-onset ataxia. No pathogenic repeat expansions compatible with diagnosis of SCA4 were identified in this extended cohort.

Potential dual genetic diagnoses and additional repeat expansions

Among the 171 patients with confirmed hereditary ataxias, 5 (2.9%) carried reduced-penetrant repeat expansion(s) compatible with a potential dual diagnosis of two distinct late-onset ataxias (Figure 2B, dashed line; Supplementary Table 5). An additional 6 patients (3.5%) carried an intermediate *FGF14* allele as a second repeat expansion (Figure 2B), while another 6 patients (3.5%) had biallelic expansions in loci associated with autosomal-dominant ataxia, including one patient each with SCA6 and SCA8, and four with SCA27B (exclusively intermediate alleles at the second allele). Furthermore, 10 patients (5.8%, not shown in Figure 2B) were also carriers for *RFC1* spectrum disorder (9 patients) or FRDA (1 patient).

Family history

Among 119 patients with autosomal-dominant repeat disorders, information on family history was available for 73 patients (Supplementary Table 5). Of these, only 42% reported having a family member affected by ataxia. When restricting the analysis to the 79 patients with fully penetrant repeat expansion, this proportion increased slightly to 55%.

Genotypes and phenotypes of repeat-associated ataxias and reassessment of carrier frequencies

ATN1 (Dentatorubral-Pallidoluysian Atrophy (DRPLA)

One patient was diagnosed with DRPLA carrying a repeat expansion in the lower range of pathogenicity (50 RU) in line with late disease onset in the early seventies. The clinical presentation included gait ataxia and oculomotor disturbances. Notably, the patient did not exhibit other common DRPLA-associated features such as myoclonus, dystonia, chorea, epilepsy or additional movement disorders, but showed mild memory impairment.

ATXN1-3, *CACNA1A* (SCA1–3, 6)

Patients diagnosed as affected by SCA1–3 and 6 carried pathogenic repeat expansions with typical compositions.Notably, one patient with SCA6 was found to be homozygous for a 20-repeat expansion in *CACNA1A*, a finding consistent with previous reports of pathogenicity at this

repeat size. Cumulative clinical features for all patients are summarized, with detailed individual data provided in Supplementary Table 6. In addition to cerebellar symptoms, peripheral neuropathy was a consistent finding across all patients diagnosed with SCA1–3 and SCA6.

ATXN8/*ATXN8OS* (SCA8)

In addition to various cerebellar signs, 66.7% of patients with a pathogenic repeat expansion in *ATXN8/ATXN8OS* frequently exhibited additional non-cerebellar features. These included both hyperkinetic and hypokinetic movement disorders, as well as upper motor neuron signs (Supplementary Table 6). This reflects the known phenotypic variability associated with SCA8.

ATXN10 (SCA10)

In two related patients (father and daughter), we identified a complex repeat expansion in *ATXN10* [$(ATTCT)_{\sim 130}(ATTCC)_{\sim 1400}$] compatible with SCA10. Unlike the pure expansions of the wildtype-motif $(ATTCT)_{exp}$ typically associated with SCA10, these patients exhibited heterogeneous expansions of the wild-type motif $(ATTCT)_{exp}$ and an alternative motif $(ATTCC)_{exp}$, previously suggested as a disease-modifier increasing risk for seizures.[37,38] Clinically, neither patient exhibited seizures, and both primarily presented with gait disturbance and other cerebellar features. EEG showed a normal result in one patient, while abortive spike-wave complexes were present in the other patient.

DAB1 (SCA37)

A complex heterozygous $[(ATTTT)_{155}(ATTTC)_{100}(ATTTT)_{10}]$ repeat expansion was identified in one female, compatible with SCA37. The repeat configuration differed slightly from previously reported patterns, which typically consist of $([(ATTTT)_{7\text{-}400}(ATTTC)_{31\text{-}79}(ATTTT)_{58\text{-}420}])$.[20,39] The pathogenic $(ATTTC)_n$ insertion was unusually long and shifted towards the 3'-end of the ATTTT expansion, leading to an atypically short 3'-ATTTT repeat. Clinically, the patient presented with a pure cerebellar ataxia phenotype, including oculomotor disturbances, limb and gait ataxia, and dysarthria. Symptom onset occurred at age 43. No additional non-cerebellar features—such as motor neuron involvement or peripheral neuropathy—were observed.

FGF14 (SCA27B)

In addition to pure $(GAA)_{>180\ RU}$ expansions, 15 patients carried *FGF14* expansions (>100 RU) involving an alternative repeat motif not yet associated with SCA27B. The most common

variant, $(GAAGGA)_{exp}$, observed in 14 patients, was often flanked by short stretches of $(GAA)_{exp}$ as previously reported.[34] One patient exhibited a novel repeat configuration, $(GGA)_{75}(GAA)_{300}$, currently classified as of unknown significance. While the pathogenic $(GAA)_{exp}$ expansion exceeds the established threshold of pathogenicity, its co-occurrence with another motif complicates interpretation.

Patients diagnosed as affected by SCA27B had a median age of onset of 62 years and most frequently presented with gait ataxia (95% of patients) and oculomotor signs (92%), while limb ataxia was only present in 66% of patients (Figure 3A, Supplementary Table 6). Approximately half of the patients exhibited an episodic onset or course (52%) and downbeat nystagmus (54%), both considerable more specific features of SCA27B. Various forms of neuropathy were present in about 43% of patients. Notably, 25% of patients for whom respective clinical data were available fulfilled the clinical triad of cerebellar ataxia, neuropathy and bilateral vestibulopathy (CANVAS). No correlation between age-of-onset and repeat length could be observed ($r < 0.01$, Figure 3C).

FMR1 (FXTAS)

Six male patients and one female patient were identified as carriers of an *FMR1* premutation, potentially associated with Fragile X-associated Tremor/Ataxia Syndrome (FXTAS). The median age at onset of cerebellar ataxia was 66 years. Based on established diagnostic criteria for FXTAS, which include both clinical features and characteristic radiological findings, three male patients met the criteria for definite FXTAS, and two were classified as probable cases. Due to incomplete clinical and imaging data, the likelihood of FXTAS in the female carrier could not be assessed.[40]

FXN (FRDA)

Among the three patients diagnosed with FRDA, two individuals—aged 38 (patient 330) and 50 (patient 263)—carried one short allele (100 and 160 repeat units, respectively) and one large expanded allele. This configuration is consistent with (very) late-onset FRDA, and both patients exhibited a milder phenotype being less specific for FRDA. Patient 330 presented with pronounced gait ataxia, dysarthria, spasticity, sensory neuropathy, and evidence of both pyramidal and dorsal column involvement based on evoked potential studies. Patient 263 showed

cerebellar ataxia with marked gait ataxia, dysarthria, and spasticity, but retained deep tendon reflexes and had no evidence of dorsal column involvement or neuropathy based on sensory evoked potentials and electroneurography. This atypical presentation delayed diagnosis for over nine years. The third patient, aged 24, carried large biallelic *FXN* repeat expansions along with an incompletely penetrant repeat expansion in *TBP*. According to the referring neurologist, this patient displayed a classical FRDA phenotype.

In addition, four individuals were identified as heterozygous carriers for FRDA. Two carried large expanded alleles (827 and 915 repeat units), typically associated with childhood- or adolescent-onset FRDA in biallelic cases. The other two carried short expansions (91 and 93 repeat units), previously linked to (very) late-onset and atypical FRDA presentations. Truncating *FXN* variants *in trans* were excluded in all of these individuals.

Five additional individuals carried complex heterozygous repeat expansions with alternative motifs: four exhibited the known $(GAAGGA)_{exp}$ configuration, and one harbored a novel $(GAAGAG)_{exp}$ motif flanked by a $(GAA)_{exp}$ segment. These were classified as variants of uncertain significance. While short $(GAA)_{exp}$ expansions interrupted by alternative motifs have been associated with very late-onset FRDA, pure $(GAAGGA)_{exp}$ expansions are currently considered likely non-pathogenic.[12,41,42] Therefore, only the four individuals with a heterozygous pure $(GAA)_{exp}$ expansion were included as FRDA carriers. Based on the 510 individuals who tested negative for biallelic *FXN* expansions, the carrier frequency for FRDA in this cohort was estimated at 0.8% (1:125).

NOP56 (SCA36)

One patient was diagnosed with SCA36 based on a large GGCCTG-expansions (>650 RU) in *NOP56*. The clinical presentation was consistent with a mild, pure cerebellar ataxia, including oculomotor abnormalities, discrete limb ataxia, and mild gait disturbance. Notably, the patient did not exhibit dysarthria or hyperreflexia—features commonly associated with SCA36 and disease onset occurred at age 45.[43] Atrophy and fasciculation of the tongue were not noticed as another common features of SCA36.[43]

RFC1 (*RFC1* spectrum disorder/CANVAS)

The majority of patients with *RFC1* spectrum disorder (40 out of 42 patients) carried biallelic pathogenic $(AAGGG)_{>250\ RU}$ repeat expansions, while the remaining two carried a heterozygous $(AAAGG)_{>500\ RU}$ expansion on one allele and a $(AAGGG)_{>250\ RU}$ expansion on the other. Characterisation of *RFC1* repeats revealed significant heterogeneity at this polymorphic locus, with 13 distinct repeat configurations identified (Figure 4, Supplementary Table 7). Notably, both interallelic and intraallelic heterogeneity were observed. The 'AAAGGG' motif was never observed as a pure expansion but was always part of a complex repeat structure containing 'AAAGG' and 'AAGGG' stretches (therefore denoted as $(AAAGGG)_{complex}$), similar to $(ACGGG)_{complex}$ expansions. One allele exhibited a unique composite configuration, $(AAAGG)_{317}(AAGGG)_{210}$, consisting of two pure repeats currently classified as of unknown significance. Sizing was possible for most *RFC1* repeat expansions; however, 40 out of 108 $(AAGGG)_{>250\ RU}$ expansions could not be precisely measured due to broken reads. Additionally, 34 patients were identified as carriers for *RFC1* spectrum disorder based on a pathogenic repeat expansion on one allele ($(AAAGG)_{>500\ RU}$ in 7 patients and $(AAGGG)_{>250\ RU}$ in 27 patients)). Of these, one patient with cerebellar ataxia and neuropathy carried a pathogenic $(AAGGG)_{556}$ expansion on one allele and an $(AAAGG)_{360}$ repeat expansion on the other, which fell below the current pathogenicity threshold. NGS analysis in all carriers excluded pathogenic variants in *trans*. Based on these 34 cases, we calculated a carrier frequency of 7.2% for *RFC1* spectrum disorder among the 471 patients in whom *RFC1* was not the primary cause of ataxia.

Patients diagnosed with *RFC1* spectrum disorder had a median age of onset of 55 years (Supplementary Table 6). Chronic cough was a common symptom, reported in 68% of patients (Figure 3B). Among those with available data, all were affected by peripheral neuropathy, which was classified as axonal based on electroneurography findings and in 18/25 patients with information available sensory and the other 7/25 patients sensorimotor. Of these, 26% presented only with neuropathy as additional symptom to cerebellar ataxia, while the remaining 74% also exhibited bilateral vestibulopathy (BVP), thereby fulfilling the diagnostic criteria for the CANVAS triad.

TBP (SCA17)

One patient was identified carrying a reduced-penetrance allele in *TBP* along with a variant of uncertain significance (VUS) in *STUB1*. The individual presented with a complex clinical phenotype, including cerebellar ataxia, seizures, and myoclonus.

NGS analysis of patients with negative or inconclusive Clin-CATS results

Of the 342 patients without a likely causative repeat expansion, 141 underwent NGS to identify pathogenic SNVs or CNVs in ataxia-related genes (Supplementary Table 5, 8). A hereditary cause of ataxia was found in 23 patients (6.7%), revealing a spectrum of ataxia syndromes, including SCA11, 13, 14, 28, 42, 48, R8, episodic ataxia type 2, *POLR3A*-associated ataxia, ataxia teleangiectasia). Additional diagnoses included hereditary spastic paraplegias with overlapping ataxia phenotypes (SPG5, 7, 39, 79A) and syndromes with cerebellar involvement, such as Tay-Sachs disease, neuropathy-ataxia-retinitis pigmentosa (NARP) syndrome, and *STXBP1*-associated developmental and epileptic encephalopathy (as somatic mosaicism).

Discussion

Diagnostic findings and prevalence of repeat-associated ataxias

The field of hereditary repeat-associated ataxias has made remarkable advancements since 2019, e.g. marked by the discovery of two additional repeat disorders that likely account for a high proportion of ataxia previously classified as sporadic.[5,7–9] These breakthroughs, driven by LRS, have not only facilitated novel disease discoveries but have also enabled long-pending revisions of clinical diagnostics of repeat disorders.[24,25] However, the wide implementation of LRS in clinical genetics remains limited. In this study, we analysed 513 patients with adult-onset ataxia using LRS-based Clin-CATS. This is one of the largest cohorts of ataxia patients evaluated since the discovery of SCA27B, to date, and the only one analysed by parallel LRS-based repeat analysis.

The diagnostic yield in this clinical cohort was high, with 33% being diagnosed with repeat-associated ataxia, increasing to 38% when intermediate *FGF14* alleles (180–249 RU) were considered. Even if causative SNVs and CNVs were not analysed in all patients, our study shows that repeat expansions are much more frequent than other genetic alterations causing ataxia (33% vs. 7%). This emphasizes the importance of prioritizing repeat analysis as an initial genetic test in patients with adult-onset ataxia, either in parallel or after the exclusion of acquired forms of ataxia.

Only 55% of patients with fully penetrant autosomal-dominant repeat expansions reported a positive family history of ataxia. In the remaining 45%, the absence of a reported family history may be explained by subtle or subclinical symptoms in affected relatives—detectable only through targeted neurological examination—or by intergenerational repeat instability. In several repeat-associated disorders, such as SCA1–3, expansions are known to increase in size during transmission, leading to more severe phenotypes in subsequent generations, a phenomenon referred to as anticipation.[44–46] Repeat expansions in *FGF14* are especially instable with substantial variations (increase and decrease) of repeat size across generations and, in some cases, shift between the pathogenic and non-pathogenic range within a single family.[47] As such, restricting diagnostic analysis to patients with a positive family history risks missing a considerable number of cases. . Consistent with findings in other populations of European ancestry, SCA27B is the most frequent adult-onset ataxia in our cohort (16.2%), followed by *RFC1* spectrum disorder (8.2 %).[48–51] The prevalence of SCA27B may be even higher considering multiple lines of evidence indicating a lower threshold for pathogenicity (20.9% when considering intermediate *FGF14* alleles).[34,52,53] Many SCA27B cases may remain undiagnosed due to subtle, slowly progressive symptoms and age-related comorbidities.[54] Other repeat disorders were substantially less common, collectively accounting for as many cases of ataxia as SCA27B alone.

Parallel repeat analysis identified rare repeat-associated ataxias, including SCA10, 36 and 37, with only SCA37 previously reported in a German patient cohort.[55–57] After updating our repeat assay by including the *ZFHX3* repeat, analysis of 347 additional ataxia patients did not reveal any pathogenic repeat expansion in this locus, indicating a low frequency of SCA4 in the population studied.

Determining the relative frequency of the individual repeat-associated ataxias is inherently challenging, regardless of the approach. The main bias in our study is the potential enrichment of patients who previously tested negative for established repeat disorders such as SCA1–3 and 6. A comparison with historical PCR-based repeat testing at our institute in the last two decades (Supplementary Figure 2) suggests SCA1–3 and 6 may be 1.9 to 6.5 times underrepresented in this cohort. Nevertheless, SCA27B and *RFC1* spectrum disease likely remain the most common repeat-associated ataxias, further supported by the high carrier frequency for *RFC1* spectrum disorder.

Phenotypic overlap of repeat-associated ataxias and correlation of age of onset and repeat length in SCA27B

Clinical characterization in our cohort highlights the significant phenotypic overlap among different repeat-associated ataxia. Differentiation based on clinical features alone would have been challenging in many cases: patients diagnosed with SCA1–3, SCA6, SCA36, and SCA37 all exhibited non-specific signs of cerebellar ataxia, occasionally accompanied by peripheral neuropathy. Without genetic testing, accurate subtype classification would have been possible only to a limited extent. Among patients diagnosed with SCA27B, 75% displayed either an episodic disease onset/course and/or downbeat nystagmus, features that have been suggested as relatively specific indicators for this condition[5,9,10,54,58] Similarly, 68% of patients with diagnosis of *RFC1* spectrum reported a chronic cough, a known presyndromal hallmark of the disease.[59] However, the presence of the CANVAS triad (cerebellar ataxia, neuropathy, and bilateral vestibulopathy) alone is insufficient to distinguish between these two entities. In our study, 25% of patients with SCA27B also fulfilled the CANVAS triad, consistent with previous reports, although the characteristics of neuropathy and vestibular involvement tend to differ between SCA27B and *RFC1* spectrum disorder.[60,61]

Current literature offers conflicting findings regarding an inverse relationship between repeat length and age of onset in SCA27B. In agreement with several previous studies, we did not observe a statistically relevant correlation in our cohort.[15,34,50,58] Nevertheless, it remains plausible that a weak inverse correlation exists and may only become apparent in larger, more age-diverse cohorts.[5,9,34]

It is important to note that clinical data in this study were collected retrospectively and not through standardized neurological assessments. The primary goal of evaluation was to determine the etiology of cerebellar ataxia, and therefore the extent and type of clinical information varied across patients. As a result, phenotypic bias—particularly for conditions represented by small numbers of cases—cannot be excluded.

Dual diagnoses and their phenotypic impact

Through parallel analysis, 17 of 171 patients with repeat-associated ataxia were found to carry an additional potentially pathogenic repeat expansion, either in the same or a different gene. The phenotypic impact of these additional repeat expansions is of particular interest. In this study, all additional expansions in other loci fell within the range of reduced penetrance or had unconfirmed pathogenicity, raising the possibility that they do not contribute to the phenotype, as has been recently described for *FGF14* expansions in patients with FRDA.[62] Biallelic pathogenic expansions in autosomal-dominant ataxias were previously reported and were also observed in this cohort. In patients with SCA6 and SCA27B, the second expansion is likely to act as a disease modifier, increasing disease severity as previously reported.[34,63,64] However, no such modifying effect could be confirmed for patients with SCA8.[65] Given the phenotypic overlap and variability in disease severity, longitudinal studies in larger cohorts, including family-based investigations, are needed to confirm the potential effects of dual diagnoses.

High carrier frequency of RFC1 spectrum disorder

The high *RFC1* carrier frequency of 7.2% for spectrum disorder is striking, corresponding to an estimated prevalence of 1:772. In principle, this high carrier frequency in a cohort of ataxia patients could indicate monoallelic pathogenic expansions in *RFC1* as a risk factor for sporadic ataxia or that a pathogenic variant *in trans* and, therefore, *RFC1* spectrum disorder may be missed in some ataxia patients. However, the carrier frequency in this cohort aligns with that observed in a large cohort of 29,496 controls with European population background (8.0%) analyzed by long-read sequencing.[66] This study can be assumed to be more representative than earlier studies, which observed lower heterozygote carrier frequencies of 0.7–6.5%.[23] As such, an enrichment of monoallelic pathogenic expansions in *RFC1* in ataxia patients not affected by

RFC1 spectrum disorder is unlikely. Compared to an estimated prevalence of sporadic adult-onset ataxias (2–12:100,000) and autosomal recessive hereditary ataxias (1-9:100,000), the estimated prevalence of *RFC1* spectrum disorder exceeds that of both conditions together by at least six-fold.[67–69] This discrepancy may reflect underdiagnosis of *RFC1* spectrum disorder due to mild or atypical presentation without ataxia (e.g. isolated neuropathy or vestibulopathy) or incomplete penetrance of biallelic pathogenic repeat expansions.[70] Longitudinal screening of older, unaffected individuals with biallelic repeat expansions could help to address this question.

Reassessing FRDA carrier frequency

For FRDA, we determined a carrier frequency of 0.8% (1:125) in our cohort, based on the assumption that short, heterogeneous alleles containing alternative motifs such as $(GAAGGA)_{exp}$ are non-pathogenic. The observed carrier frequency for FRDA in our cohort is lower than reported estimates (1:60 – 1:100).[71]

The most likely explanation for this discrepancy is statistical variation due to our cohort size. However, we also hypothesize that earlier studies may have overestimated carrier frequencies due to limitations of PCR-based genotyping methods, which cannot reliably distinguish pathogenic pure (GAA)exp alleles from non-pathogenic expansions with alternative repeat motifs. Consequently, inclusion of such non-pathogenic alleles in previous analyses may have artificially inflated carrier frequency estimates.

To assess whether heterozygous carrier frequencies for FRDA can be directly translated into disease prevalence, future studies should evaluate the clinical relevance of biallelic short expansions (<100 repeat units). In our cohort, such short alleles were found among carriers at similar frequencies as large expansions. It remains to be determined whether these biallelic short expansions are associated with very late-onset FRDA in only a small subset of individuals, or whether they exhibit high penetrance more broadly.

Challenges in the assessment of repeat expansions

Advances in LRS-based repeat analysis highlight unresolved challenges in assessing repeat expansions, particularly in determining their pathogenicity, penetrance, and the prediction of age-of-onset, severity and phenotypic presentation. These hurdles are pronounced for recently

described conditions like *RFC1* spectrum disorder and SCA27B, where a significant number of diagnoses remain inconclusive. For SCA27B, a major issue is the uncertain threshold of pathogenicity. While current studies set the threshold for pathogenicity with reduced penetrance at 250 RU, other studies suggest lower values (180 or 200 RU).[15,34,52,60] Establishing a definitive threshold of pathogenicity is challenging due to the extensive mosaicism of repeat expansions in *FGF14*. Additionally, inherent biases in both PCR-based methods and amplification-free long-read sequencing techniques can lead to slight variations in the median repeat length (Supplementary Figure 1). Consequently, studies aimed at determining length thresholds must carefully benchmark their methods to account for these potential discrepancies. In this study, 35 patients carried expansions in the 180-249 RU range, which require further validation regarding an SCA27B diagnosis. Additionally, novel *FGF14* repeat configurations, such as $(GGA)_{75}(GAA)_{300}$, observed in this study, which carry large continuous stretches of the pathogenic motif, require further investigation.

In *RFC1* spectrum disorder, up to 24 distinct repeat motifs have been identified (confirmed 13 motives in Supplementary Table 1), some extremely rare or highly heterogenous, complicating pathogenicity assessment.[19,22,24,25,72] In this study, 24 alleles (2.3%) carried repeat configurations of unknown significance, such as ($(ACGGG)_{complex}$, $(AACGG)_{exp}$, $(ACAAG)_{exp}$ and $(AGGGG)_{exp}$). Most patients carried a non-pathogenic repeat configuration on the second allele, largely excluding *RFC1* spectrum disorder. However, one individual carried repeat configurations of unknown significance on both alleles ($(ACGGG)_{complex}/(AACGG)_{exp}$), hindering the exclusion of *RFC1* spectrum disorder. Based on the assumption that GC-content and repeat length determine pathogenicity, we speculate that the 80 RU 'AACGG' expansion is non-pathogenic, as it has the same GC-content as the pathogenic 'AAGGG' motif but is shorter than the current pathogenicity threshold (250 RU).[19] Similarly to *FGF14*, LRS deciphered composite expansions in *RFC1*, such as $(AAAGG)_{317}(AAGGG)_{210}$, which require an assessment of pathogenicity. As the formation of stable G-quadruplex structures has been associated with the pathogenic 'AAGGG' motif, analysing secondary RNA/DNA structures across all known motifs could contribute to the assessment of their clinical significance in conjunction with further elucidating the pathomechanism of *RFC1* spectrum disorder.[73–75] Adding to the complexity of *RFC1* repeat expansion assessment, two motifs ('AAGGG' and 'AAAGG'), previously assumed as solely pathogenic or non-pathogenic, respectively, are now known to be pathogenic depending

on their length.[19] Current thresholds of pathogenicity require further confirmation and refinement, e.g. through large (meta)analyses.

A significant proportion of ataxia patients with likely causative repeat expansion, 44% (75 out of 171 patients) in this study, carry reduced-penetrance expansions in loci such as *ATXN8/ATXN8OS*, *TBP*, *FGF14*, and *FMR1*. Emerging data indicate that not only incomplete penetrance but also repeat expansions with currently assumed full penetrance are frequent in healthy individuals.[66,76] Genetic findings must, therefore, be critically evaluated in the context of clinical findings, phenotype and differential diagnoses as penetrance might be lower than currently assumed, leading to a potentially unrelated coincidence of adult-onset ataxia and a repeat expansion associated with ataxia.

Given the phenotypic overlap in patients with adult-onset ataxia hindering clinical classification, molecular genetic diagnoses should be refined and may benefit from the inclusion of additional genetic parameters. Depending on the underlying pathomechanism, the prediction of penetrance and disease severity may be improved by analysing flanking regions, repeat interruptions, alternative repeat motifs, somatic mosaicism, and methylation. For example, in case of Huntington's disease, somatic instability contributes for variability in penetrance, age of onset, and clinical severity, factors that extend beyond repeat length alone.[77–79] The same might apply to certain SCAs, as indicated by a recent study.[80] The degree of somatic instability may be inferred from repeat length heterogeneity in blood, as accessible through Clin-CATS. Similarly, indirect markers, such as repeat interruptions and DNA repair gene variants, may further indicate repeat stability.[81,82] For repeat disorders driven by loss-of-function mechanisms, such as FRDA and likely SCA27B, methylation patterns of regulatory regions could serve as markers of residual gene expression, potentially predicting penetrance and clinical severity.[83,84] In addition to investigating additional parameters, pathogenic variants in other genes might impact or determine the phenotype, as described for SCA48, where intermediate alleles in *TBP* likely act as disease modifiers.[36]

Targeted LRS as a superior tool for repeat analysis and remaining technological challenges

While NGS has fundamentally changed the detection of SNVs, methods for repeat analysis have stagnated for decades. The limitations of traditional PCR-based approaches are particularly evident in complex repeat disorders such as *RFC1* spectrum disorder. Even combining several PCR-based analyses could not reliably determine disease and carrier status for all patients analysed in this study due to the inability to measure expansions larger than 150–200 RU and the restriction of the screening to a limited set of repeat motifs. Recent bioinformatic methods, such as ExpansionHunter, offer a valuable tool for screening repeat expansions in short-read genomes.[85] However, these methods cannot accurately determine the repeat length of larger expansions and repeat interruptions due to read length limitations. Moreover, while their sensitivity and specificity are high, they remain lower compared to PCR-based methods, resulting in a significant false-positive rate, which necessitates confirmation by complementary methods.[86,87]

LRS has the potential to overcome persisting limitations while enabling the incorporation of additional genetic features, such as repeat interruptions and methylation. With decreasing costs and increasing sequencing accuracy, long-read genomes will likely become the gold standard in the diagnosis of hereditary ataxias, as repeat-associated ataxias can be detected in parallel to other genetic variants. However, current targeted approaches are required for cost-efficiency and precision. Current LRS platforms include ONT or Pacific Biosciences, with several targeted approaches including CRISPR/Cas9-based enrichment, and ONT bioinformatic-based enrichment with adaptive sampling.[88] However, a comprehensive benchmarking, comparison, and clinical validation of all LRS approaches is currently lacking. Individual studies highlight specific limitations, such as repeat expansions in *FXN* or *RFC1* carrying specific motifs missed in PacBio genome sequencing or missed expansions in adaptive sampling due to low coverage.[86,89]

In contrast, Clin-CATS, based on CRISPR/Cas9 target enrichment, ensures high coverage of all repeat loci and replaces at least 18 individual PCR analyses with a single, more comprehensive test. However, sizing very large expansions, such as the pathogenic motif ‘AAGGG’ in *RFC1*, remains a challenge. While Southern blotting has detected expansions as large as 3885 RU, no

'AAGGG'-repeat expansion larger than 1395 RU was detected in our study.[90] This could be explained by the fragmentation of large repeat expansions composed of the pathogenic 'AAGGG' motif, potentially as a consequence of its tendency to form G-quadruplex DNA structures in combination with ONT's bias toward shorter fragments. As the presence of expansions above the threshold of pathogenicity is reliably detected also, in these cases, the diagnostic precision of Clin-CATS is not affected by this limitation. Further optimisation could extend the upper limit of detectable expansions, but complementary methods such as Southern blotting or Optical Genome Mapping may still be required for extremely large expansions.[91] Finally, to maintain the robust performance of Clin-CATS in routine diagnostics, we adapted the workflow to be compatible with the new R10 flow cell chemistry following the discontinuation of the R9 flow cell in a recent work.[92] Improvements to the library preparation protocol and bioinformatics pipeline have ensured the continuation of Clin-CATS with the latest technology.

Conclusion

This study found and confirmed the high prevalence of repeat disorders as the cause of adult-onset ataxia, with SCA27B and *RFC1* spectrum disorder as the most common conditions. We identified several novel genotypes and variations of the known repeat configurations, further illustrating the genetic heterogeneity of repeat-associated ataxias. Our study also highlights the methodological advantages of LRS-based repeat analysis with Clin-CATS, showing a stable and robust performance and cost-efficiency. Given the demonstrated superiority of parallel analysis with LRS, we strongly advocate for its adoption as first-tier analysis in patients with adult-onset ataxia before performing NGS exome or genome analysis.

Data availability

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

Acknowledgements

T.K. acknowledges support by the European Union, project European Rare Disease Research Alliance (ERDERA, GA No 101156595), funded under call HORIZON-HLTH-2023-DISEASE-07, and support by the German Federal Ministry of Research, Technology and Space (BMFTR, Bonn, Germany) through grants to the German Network for Mitochondrial Disorders (mitoNET, 01GM1906A) and to the E-Rare project GENOMIT (01GM1920B). T.K. is a member of the European Reference Network for Rare Neurological Diseases (ERN-RND).

Funding

AZ received funding from the Federal Ministry of Research and Education (BMBF, 01 EO 1401) and HORIZON 2020 MSCA-DN PROVIDE supporting this work.

Competing interests

CS-G received speakers' and/or consultants' honoraria from Amicus Therapeutics, Argenx, Alexion, Amgen,Hormosan Pharma, Immunovant, Janssen, Lupin Pharmaceuticals, Novartis, Sanofi-Genzyme, Roche and UCB.

MCL has received travel and accommodation expenses to speak at ONT conferences.

MS received speaker fees from Abbott, Auris Medical, Biogen, Eisai, Grünenthal, GSK, Henning Pharma, Interacoustics, J&J, MSD, NeuroUpdate, Otometrics, Pierre-Fabre, TEVA, UCB, and Viatris. He receives support for clinical studies from Decibel, U.S.A., Cure within Reach, U.S.A. and Heel, Germany. He distributes "M-glasses" and "Positional vertigo App." He acts as a consultant for Abbott, AurisMedical, Bulbitec, Heel, IntraBio, Sensorion, Vifor and Vertify. He is investor, patent-and share-holder of IntraBio.

JL received speaker fees from Bayer Vital, Biogen, and Roche; consulting fees from Axon Neuroscience and Biogen; author fees from Thieme medical publishers and W. Kohlhammer GmbH medical publishers; non-financial support from Abbvie; and compensation for duty as part-time CMO from MODAG, all outside the submitted work.

TK received received research support and/or personal compensation from Santhera Pharmaceuticals, Chiesi GmbH, and GenSight Biologics.

FS received honoraria for advisory boards of Amylyx, Alnylam, and Alexion and publication royalties from W. Kohlhammer GmbH medical publishers.

AZ received speaker fees from Dr. Willmar Schwabe GmbH, AstraZeneca, Pfizer, research support from Dr. Willmar Schwabe GmbH and Cures Within Reach. Received publication royalties from Academic Press, W. Kohlhammer GmbH, Springer and Thieme medical publishers.

The other authors declare no conflict of interest regarding this work.

The authors declare that the submitted work was carried out in the absence of any professional or financial relationships that could potentially be construed as a conflict of interest.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Jayadev S, Bird TD. Hereditary ataxias: overview. *Genet Med*. 2013;15:673-683.
2. Hadjivassiliou M, Martindale J, Shanmugarajah P, et al. Causes of progressive cerebellar ataxia: prospective evaluation of 1500 patients. *J Neurol Neurosurg Psychiatry*. 2017;88:301-309.
3. de Silva RN, Vallortigara J, Greenfield J, Hunt B, Giunti P, Hadjivassiliou M. Diagnosis and management of progressive ataxia in adults. *Pract Neurol*. 2019;19:196.
4. Chintalaphani SR, Pineda SS, Deveson IW, Kumar KR. An update on the neurological short tandem repeat expansion disorders and the emergence of long-read sequencing diagnostics. *Acta Neuropathol Commun*. 2021;9:98.

5. Pellerin D, Danzi MC, Wilke C, et al. Deep Intronic FGF14 GAA Repeat Expansion in Late-Onset Cerebellar Ataxia. *N Engl J Med*. 2022;388:128-141.

6. Wallenius J, Kafantari E, Jhaveri E, et al. Exonic trinucleotide repeat expansions in ZFHX3 cause spinocerebellar ataxia type 4: A poly-glycine disease. *Am J Hum Genet*. 2024;111:82-95.

7. 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.

8. 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.

9. Rafehi H, Read J, Szmulewicz DJ, et al. An intronic GAA repeat expansion in FGF14 causes the autosomal-dominant adult-onset ataxia SCA50/ATX-FGF14. *Am J Hum Genet*. 2023;110:105-119.

10. Wirth T, Clément G, Delvallée C, et al. Natural History and Phenotypic Spectrum of GAA-FGF14 Sporadic Late-Onset Cerebellar Ataxia (SCA27B). *Mov Disord*. 2023;38:1950-1956.

11. Cortese A, Tozza S, Yau WY, et al. Cerebellar ataxia, neuropathy, vestibular areflexia syndrome due to RFC1 repeat expansion. *Brain*. 2020;143:480-490.

12. Currò R, Salvalaggio A, Tozza S, et al. RFC1 expansions are a common cause of idiopathic sensory neuropathy. *Brain*. 2021;144:1542-1550.

13. Gisatulin M, Dobricic V, Zühlke C, et al. Clinical spectrum of the pentanucleotide repeat expansion in the RFC1 gene in ataxia syndromes. *Neurology*. 2020;95:e2912.

14. Traschütz A, Heindl F, Bilal M, et al. Frequency and Phenotype of RFC1 Repeat Expansions in Bilateral Vestibulopathy. *Neurology*. 2023;101:e1001-e1013.

15. Méreaux JL, Davoine CS, Pellerin D, et al. Clinical and genetic keys to cerebellar ataxia due to FGF14 GAA expansions. *EBioMedicine*. 2024;99:104931.

16. Pellerin D, Wilke C, Traschütz A, et al. Intronic FGF14 GAA repeat expansions are a common cause of ataxia syndromes with neuropathy and bilateral vestibulopathy. *J Neurol Neurosurg Psychiatry*. 2024;95:175-179.

17. Strupp M, Schüler O, Krafczyk S, et al. Treatment of downbeat nystagmus with 3,4-diaminopyridine: A placebo-controlled study. *Neurology*. 2003;61:165-170.

18. Wallace SE, Bird TD. Molecular genetic testing for hereditary ataxia: What every neurologist should know. *Neurol Clin Pract*. 2018;8:27-32.

19. Dominik N, Magri S, Currò R, et al. Normal and pathogenic variation of RFC1 repeat expansions: implications for clinical diagnosis. *Brain*. 2023;146:5060-5069.

20. Seixas AI, Loureiro JR, Costa C, et al. A Pentanucleotide ATTTC Repeat Insertion in the Non-coding Region of DAB1, Mapping to SCA37, Causes Spinocerebellar Ataxia. *Am J Hum Genet*. 2017;101:87-103.

21. Ishikawa K. Spinocerebellar ataxia type 31 (SCA31). *J Hum Genet*. 2023;68:153-156.

22. Scriba CK, Stevanovski I, Chintalaphani SR, et al. RFC1 in an Australasian neurological disease cohort: extending the genetic heterogeneity and implications for diagnostics. *Brain Commun*. 2023;5:fcad208.

23. Davies K, Szmulewicz DJ, Corben LA, Delatycki M, Lockhart PJ. RFC1-Related Disease. *Neurol Genet*. 2022;8:e200016.

24. Stevanovski Igor, Chintalaphani Sanjog R., Gamaarachchi Hasindu, et al. Comprehensive genetic diagnosis of tandem repeat expansion disorders with programmable targeted nanopore sequencing. *Sci Adv*. 8:eabm5386.

25. Erdmann H, Schöberl F, Giurgiu M, et al. Parallel in-depth analysis of repeat expansions in ataxia patients by long-read sequencing. *Brain*. 2023;146:1831-1843.

26. Tanudisastro HA, Deveson IW, Dashnow H, MacArthur DG. Sequencing and characterizing short tandem repeats in the human genome. *Nat Rev Genet*. 2024;25:460-475.

27. Figueroa KP, Gross C, Buena-Atienza E, et al. A GGC-repeat expansion in ZFHX3 encoding polyglycine causes spinocerebellar ataxia type 4 and impairs autophagy. *Nat Genet*. 2024;56:1080-1089.

28. Wick RR, Judd LM, Holt KE. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol*. 2019;20:129.

29. 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.

30. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34:3094-3100.

31. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics*. 2018;34:2666-2669.

32. Giesselmann P, Brändl B, Raimondeau E, et al. Analysis of short tandem repeat expansions and their methylation state with nanopore sequencing. *Nat Biotechnol*. 2019;37:1478-1481.

33. GitHub. GitHub. February 2, 2022. Accessed February 2, 2022. https://github.com/erdogant/findpeaks

34. Mohren L, Erdlenbruch F, Leitão E, et al. Identification and characterisation of pathogenic and non-pathogenic FGF14 repeat expansions. *Nat Commun*. 2024;15:7665.

35. 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;17:405-423.

36. Barbier M, Davoine CS, Petit E, et al. Intermediate repeat expansions of TBP and STUB1: Genetic modifier or pure digenic inheritance in spinocerebellar ataxias? *Genet Med*. 2023;25:100327.

37. Morato Torres CA, Zafar F, Tsai YC, et al. ATTCT and ATTCC repeat expansions in the ATXN10 gene affect disease penetrance of spinocerebellar ataxia type 10. *HGG Adv*. 2022;3:100137.

38. Schüle B, McFarland KN, Lee K, et al. Parkinson's disease associated with pure ATXN10 repeat expansion. *NPJ Park Dis*. 2017;3:27.

39. Sanchez-Flores M, Corral-Juan M, Gasch-Navalón E, Cirillo D, Sanchez I, Matilla-Dueñas A. Novel genotype-phenotype correlations, differential cerebellar allele-specific methylation, and a common origin of the (ATTTC)n insertion in spinocerebellar ataxia type 37. *Hum Genet*. 2024;143:211-232.

40. Hunter JE, Berry-Kravis E, Hipp H, Todd PK. FMR1 Disorders. GeneReviews®. September 8, 2024. Accessed September 8, 2024. https://www.ncbi.nlm.nih.gov/books/NBK1384/

41. Stolle CA, Frackelton EC, McCallum J, et al. Novel, complex interruptions of the GAA repeat in small, expanded alleles of two affected siblings with late-onset Friedreich ataxia. *Mov Disord*. 2008;23:1303-1306.

42. Ohshima K, Sakamoto N, Labuda M, et al. A nonpathogenic GAAGGA repeat in the Friedreich gene: implications for pathogenesis. *Neurology*. 1999;53:1854-1857.

43. Ikeda Y, Ohta Y, Kobayashi H, et al. Clinical features of SCA36: a novel spinocerebellar ataxia with motor neuron involvement (Asidan). *Neurology*. 2012;79:333-341.

44. Opal, Puneet, Ashizawa, Tetsuo. Spinocerebellar Ataxia Type 1. GeneReviews®. June 7, 2021. Accessed June 7, 2021. https://www.ncbi.nlm.nih.gov/books/NBK1184/

45. Pulst, Stefan. Spinocerebellar Ataxia Type 2. GeneReviews®. June 7, 2021. Accessed June 7, 2021. https://www.ncbi.nlm.nih.gov/books/NBK1275/

46. Paulson, Henry S Vikram Paulson, Henry. Spinocerebellar Ataxia Type 3. GeneReviews®. June 7, 2021. Accessed June 7, 2021. https://www.ncbi.nlm.nih.gov/books/NBK1196/

47. Pellerin D, Del Gobbo GF, Couse M, et al. A common flanking variant is associated with enhanced stability of the FGF14-SCA27B repeat locus. *Nat Genet*. 2024;56:1366-1370.

48. Koutsis G, Kartanou C, Kontogeorgiou Z, et al. Screening for SCA27B, CANVAS and other repeat expansion disorders in Greek patients with late-onset cerebellar ataxia suggests a need to update current diagnostic algorithms. *J Neurol Sci*. 2024;467:123309.

49. Satolli S, Rossi S, Vegezzi E, et al. Spinocerebellar ataxia 27B: a frequent and slowly progressive autosomal-dominant cerebellar ataxia—experience from an Italian cohort. *J Neurol*. 2024;271:5478–5488.

50. Iruzubieta P, Pellerin D, Bergareche A, et al. Frequency and phenotypic spectrum of spinocerebellar ataxia 27B and other genetic ataxias in a Spanish cohort of late-onset cerebellar ataxia. *Eur J Neurol*. 2023;30:3828-3833.

51. Milovanović A, Dragaševic-Mišković N, Thomsen M, et al. RFC1 and FGF14 Repeat Expansions in Serbian Patients with Cerebellar Ataxia. *Mov Disord Clin Pract*. 2024;11:626-633.

52. Miyatake S, Doi H, Yaguchi H, et al. Complete nanopore repeat sequencing of SCA27B (GAA-FGF14 ataxia) in Japanese. *J Neurol Neurosurg Psychiatry*. 2024;95:1187–1195.

53. Pellerin D, Heindl F, Wilke C, et al. GAA-FGF14 disease: defining its frequency, molecular basis, and 4-aminopyridine response in a large downbeat nystagmus cohort. *EBioMedicine*. 2024;102:105076.

54. Abou Chaar W, Eranki AN, Stevens HA, et al. Clinical, Radiological and Pathological Features of a Large American Cohort of Spinocerebellar Ataxia (SCA27B). *Ann Neurol*. 2024;96:1092-1103.

55. Hengel H, Pellerin D, Wilke C, et al. As Frequent as Polyglutamine Spinocerebellar Ataxias: SCA27B in a Large German Autosomal Dominant Ataxia Cohort. *Mov Disord*. 2023;38:1557-1558.

56. Aydin G, Dekomien G, Hoffjan S, Gerding WM, Epplen JT, Arning L. Frequency of SCA8, SCA10, SCA12, SCA36, FXTAS and C9orf72 repeat expansions in SCA patients negative for the most common SCA subtypes. *BMC Neurol*. 2018;18:3.

57. Rosenbohm A, Pott H, Thomsen M, et al. Familial Cerebellar Ataxia and Amyotrophic Lateral Sclerosis/Frontotemporal Dementia with DAB1 and C9ORF72 Repeat Expansions: An 18-Year Study. *Mov Disord*. 2022;37:2427-2439.

58. Wilke C, Pellerin D, Mengel D, et al. GAA-FGF14 ataxia (SCA27B): phenotypic profile, natural history progression and 4-aminopyridine treatment response. *Brain.* 2023;146:4144-4157.

59. Cortese A, Reilly, Mary M, Houlden, Henry. RFC1 CANVAS / Spectrum Disorder. GeneReviews®. April 28, 2021. Accessed April 28, 2021. https://www.ncbi.nlm.nih.gov/books/NBK564656/

60. Pellerin D, Wilke C, Traschütz A, et al. Intronic FGF14 GAA repeat expansions are a common cause of ataxia syndromes with neuropathy and bilateral vestibulopathy. *J Neurol Neurosurg Psychiatry*. 2024;95:175-179.

61. Borsche M, Thomsen M, Szmulewicz DJ, et al. Bilateral vestibulopathy in RFC1-positive CANVAS is distinctly different compared to FGF14-linked spinocerebellar ataxia 27B. *J Neurol*. 2024;271:1023-1027.

62. Gerhart BJ, Pellerin D, Danzi MC, et al. Assessment of the Clinical Interactions of GAA Repeat Expansions in FGF14 and FXN. *Neurol Genet*. 2024;10:e200210.

63. Kato T, Tanaka F, Yamamoto M, et al. Sisters homozygous for the spinocerebellar ataxia type 6 (SCA6)/CACNA1A gene associated with different clinical phenotypes. *Clin Genet*. 2000;58:69-73.

64. Hatano Y, Ishihara T, Hirokawa S, et al. Redefining the Pathogenic CAG Repeat Units Threshold in CACNA1A for Spinocerebellar Ataxia Type 6. *Neurol Genet*. 2025;11:e200245.

65. Day JW, Schut LJ, Moseley ML, Durand AC, Ranum LPW. Spinocerebellar ataxia type 8: Clinical features in a large family. *Neurology*. 2000;55:649-657.

66. Ibañez K, Jadhav B, Zanovello M, et al. Increased frequency of repeat expansion mutations across different populations. *Nat Med*. 2024;30:3357-3368.

67. Palau F, Espinós C. Autosomal recessive cerebellar ataxias. *Orphanet J Rare Dis*. 2006;1:47.

68. Leone M, Bottacchi E, D'Alessandro G, Kustermann S. Hereditary ataxias and paraplegias in Valle d'Aosta, Italy: a study of prevalence and disability. *Acta Neurol Scand*. 1995;91:183-187.

69. Tsuji S, Onodera O, Goto J, Nishizawa M, Study Group on Ataxic Diseases. Sporadic ataxias in Japan--a population-based epidemiological study. *Cerebellum*. 2008;7:189-197.

70. Traschütz A, Cortese A, Reich S, et al. Natural History, Phenotypic Spectrum, and Discriminative Features of Multisystemic RFC1 Disease. *Neurology*. 2021;96:e1369-e1382.

71. Bidichandani, Sanjay, Delatycki, Martin B. Friedreich Ataxia. GeneReviews®. June 7, 2021. Accessed June 7, 2021. https://www.ncbi.nlm.nih.gov/books/NBK1140/

72. Schaub A, Erdmann H, Scholz V, et al. 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. *J Neurol*. 2024;271:5804-5812.

73. Abdi MH, Zamiri B, Pazuki G, Sardari S, Pearson CE. Pathogenic CANVAS-causing but not nonpathogenic RFC1 DNA/RNA repeat motifs form quadruplex or triplex structures. *J Biol Chem*. 2023;299:105202.

74. Wang Y, Wang J, Yan Z, et al. Structural investigation of pathogenic RFC1 AAGGG pentanucleotide repeats reveals a role of G-quadruplex in dysregulated gene expression in CANVAS. *Nucleic Acids Res*. 2024;52:2698-2710.

75. Singh K, Shukla S, Shankar U, et al. Elucidating the pathobiology of Cerebellar Ataxia with Neuropathy and Vestibular Areflexia Syndrome (CANVAS) with its expanded RNA structure formation and proteinopathy. *Sci Rep*. 2024;14:28054.

76. Kay C, Collins JA, Miedzybrodzka Z, et al. Huntington disease reduced penetrance alleles occur at high frequency in the general population. *Neurology*. 2016;87:282-288.

77. Swami M, Hendricks AE, Gillis T, et al. Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. *Hum Mol Genet*. 2009;18:3039-3047.

78. Ciosi M, Maxwell A, Cumming SA, et al. A genetic association study of glutamine-encoding DNA sequence structures, somatic CAG expansion, and DNA repair gene variants, with Huntington disease clinical outcomes. *EBioMedicine*. 2019;48:568-580.

79. Handsaker RE, Kashin S, Reed NM, et al. Long somatic DNA-repeat expansion drives neurodegeneration in Huntington's disease. *Cell*. 2025;188:623-639.e19.

80. Kacher R, Lejeune FX, David I, et al. CAG repeat mosaicism is gene specific in spinocerebellar ataxias. *Am J Hum Genet*. 2024;111:913-926.

81. Rajan-Babu IS, Peng JJ, Chiu R, et al. Genome-wide sequencing as a first-tier screening test for short tandem repeat expansions. *Genome Med*. 2021;13:126.

82. Bettencourt C, Hensman-Moss D, Flower M, et al. DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Ann Neurol*. 2016;79:983-990.

83. Soragni E, Herman D, Dent SYR, Gottesfeld JM, Wells RD, Napierala M. Long intronic GAA*TTC repeats induce epigenetic changes and reporter gene silencing in a molecular model of Friedreich ataxia. *Nucleic Acids Res*. 2008;36:6056-6065.

84. Lam C, Gilliam KM, Rodden LN, Schadt KA, Lynch DR, Bidichandani S. FXN gene methylation determines carrier status in Friedreich ataxia. *J Med Genet*. 2023;60:797-800.

85. Dolzhenko E, Deshpande V, Schlesinger F, et al. ExpansionHunter: a sequence-graph-based tool to analyze variation in short tandem repeat regions. *Bioinformatics*. 2019;35:4754-4756.

86. Rafehi H, Fearnley LG, Read J, et al. A prospective trial comparing programmable targeted long-read sequencing and short-read genome sequencing for genetic diagnosis of cerebellar ataxia. *Genome Res*. 2025;35:769-785.

87. Ibañez K, Polke J, Hagelstrom RT, et al. Whole genome sequencing for the diagnosis of neurological repeat expansion disorders in the UK: a retrospective diagnostic accuracy and prospective clinical validation study. *Lancet Neurol*. 2022;21:234-245.

88. Leitão E, Schröder C, Depienne C. Identification and characterization of repeat expansions in neurological disorders: Methodologies, tools, and strategies. *Rev Neurol (Paris)*. 2024;180:383-392.

89. Höps W, Weiss MM, Derks R, et al. HiFi long-read genomes for difficult-to-detect, clinically relevant variants. *Am J Hum Genet*. 2025;112:450-456.

90. Currò R, Dominik N, Facchini S, et al. Role of the repeat expansion size in predicting age of onset and severity in RFC1 disease. *Brain*. 2024;147:1887-1898.

91. Facchini S, Dominik N, Manini A, et al. Optical Genome Mapping Enables Detection and Accurate Sizing of RFC1 Repeat Expansions. *Biomolecules*. 2023;13:1546.

92. Scholz V, Schönrock V, Erdmann H, et al. Parallel in-depth analysis of repeat expansions: an updated Clin-CATS workflow for nanopore R10 flow cells. *bioRxiv*. Published online January 1, 2024:2024.11.05.622099. doi:10.1101/2024.11.05.622099

Figure legends

Figure 1 Overview of repeat-associated ataxias. Repeat-associated ataxias and their location in the genome, repeat motifs and associated genes. For all polymorphic repeat loci, the most frequent pathogenic repeat motif or repeat configuration is given. *Recently included within the last update of Clin-CATS.

Figure 2 Diagnostic results within a cohort of 513 patients with adult-onset ataxia. (A) Classification of the 513 patients analysed by Clin-CATS (all patients) and NGS (185 patients). [+] NGS testing was performed in only 118 of these patients. (B) Diagnostic results based on Clin-CATS of 513 patients with adult-onset ataxia. Flow diagram visualising diagnostic result (positive/negative), primary (dark green lines) and dual diagnoses. [*]indicates reduced penetrance; [**]patient carried additionally a VUS in *STUB1*; dashed lines indicate dual diagnoses; [±]indicates intermediate alleles in *FGF14* with unknown clinical significance. One patient with intermediate *FGF14* allele was diagnosed with SPG7 by NGS.

Figure 3 Clinical features in patients living with SCA27B and *RFC1* spectrum disorder. Frequency of different symptoms in patients living with (A) SCA27B and (B) *RFC1* spectrum disorder. (C) Correlation of age of onset and *FGF14* repeat length.

Figure 4 Repeat configurations of *RFC1*, their frequency and current clinical classification. Total number of alleles: 1026.

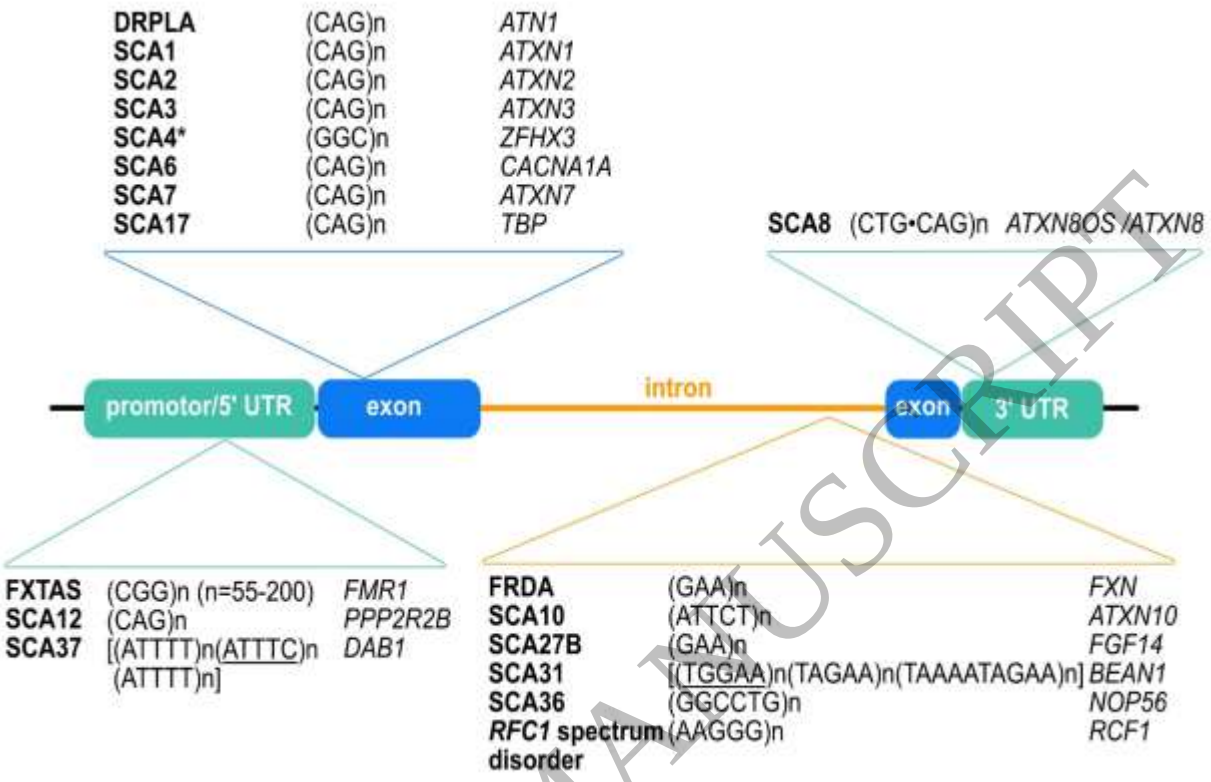

Figure 1
356x194 mm (x DPI)

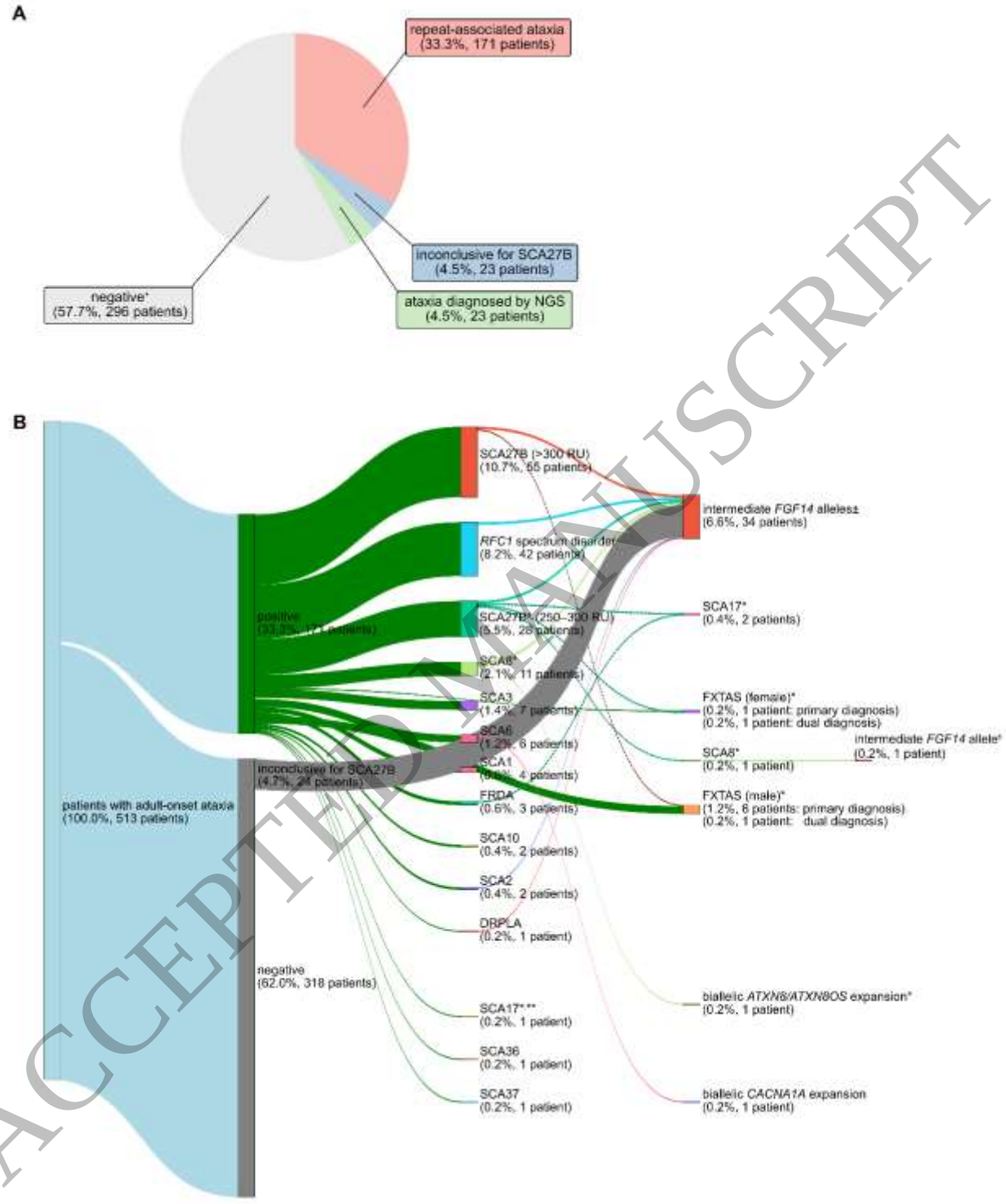

Figure 2
408x496 mm (x DPI)

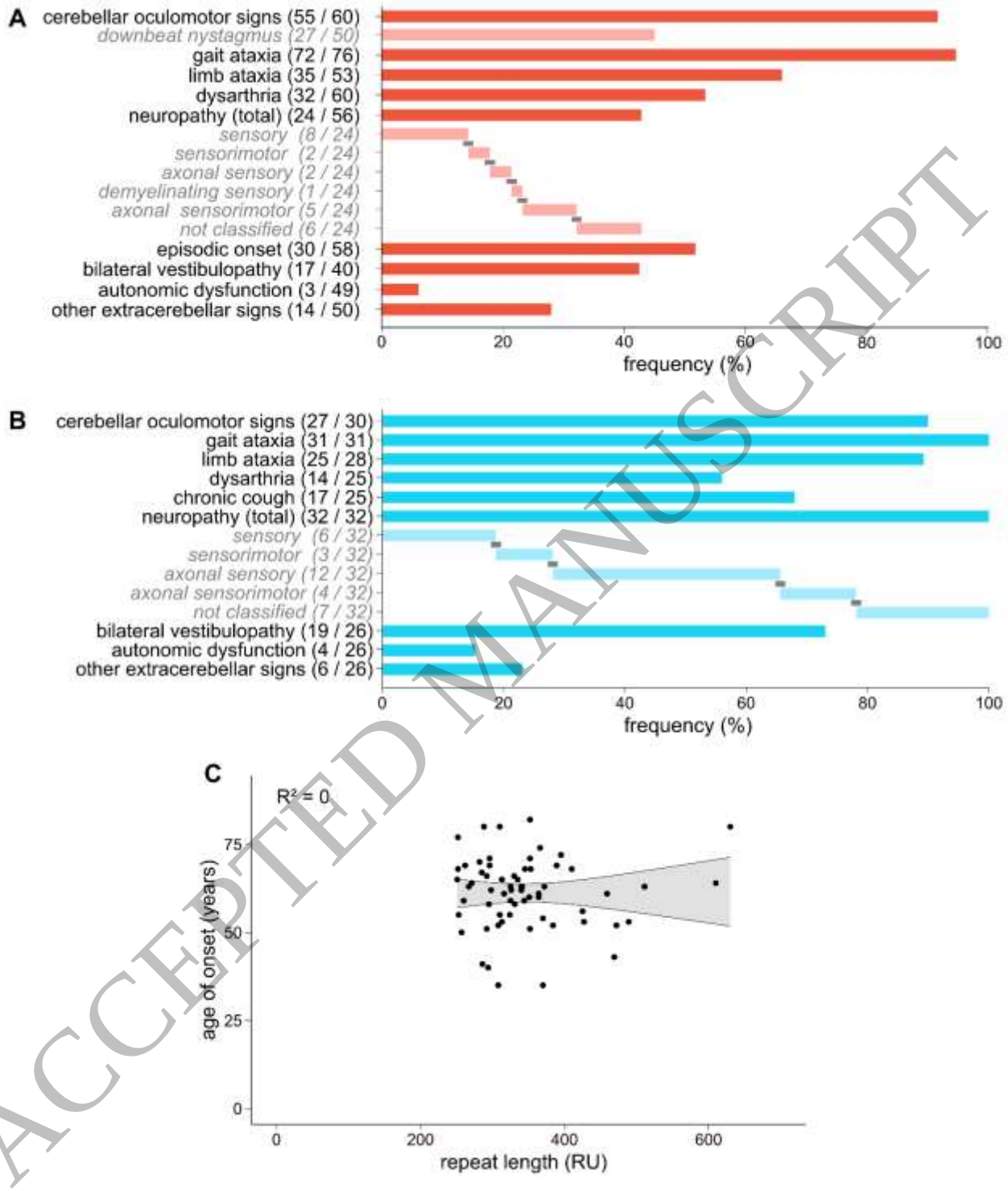

Figure 3
340x400 mm (x DPI)

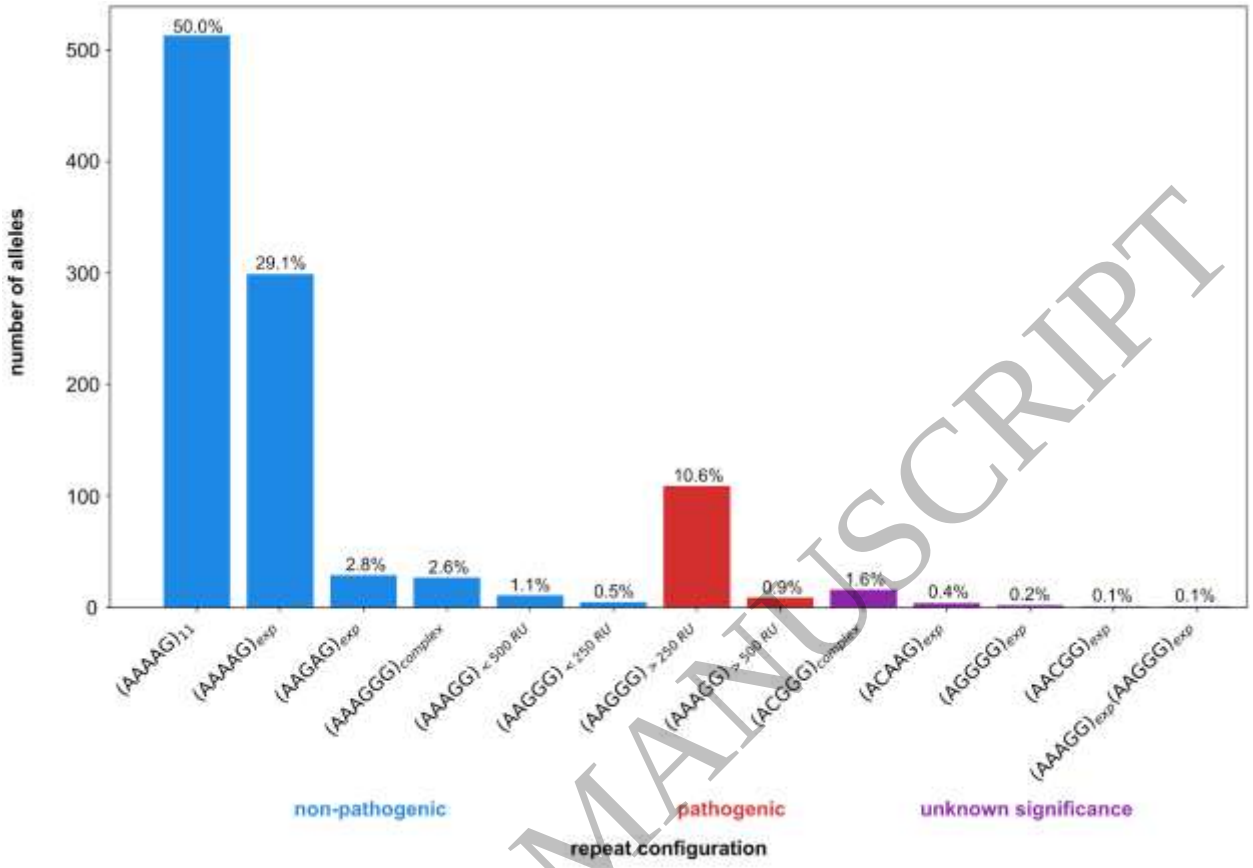

Figure 4
559x395 mm (x DPI)