Neurological disorders caused by novel non-coding repeat expansions: clinical features and differential diagnosis

Elisa Vegezzi, Hiroyuki Ishiura, D Cristopher Bragg, David Pellerin, Francesca Magrinelli, Riccardo Currò, Stefano Facchini, Arianna Tucci, John Hardy, Nutan Sharma, Matt C Danzi, Stephan Zuchner, Bernard Brais, Mary M Reilly, Shoji Tsuji, Henry Houlden, Andrea Cortese

Nucleotide repeat expansions in the human genome are a well-known cause of neurological disease. In the past decade, advances in DNA sequencing technologies have led to a better understanding of the role of non-coding DNA, that is, the DNA that is not transcribed into proteins. These techniques have also enabled the identification of pathogenic non-coding repeat expansions that cause neurological disorders. Mounting evidence shows that adult patients with familial or sporadic presentations of epilepsy, cognitive dysfunction, myopathy, neuropathy, ataxia, or movement disorders can be carriers of non-coding repeat expansions. The description of the clinical, epidemiological, and molecular features of these recently identified non-coding repeat expansion disorders should guide clinicians in the diagnosis and management of these patients, and help in the genetic counselling for patients and their families.

Lancet Neurol 2024; 23: 725–39

IRCCS Mondino Foundation, Pavia, Italy (E Vegezzi MD, S Facchini PhD); **Department of Neurology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan** (Prof H Ishiura MD, Prof S Tsuji MD); **Department of Neuromuscular Diseases** (D Pellerin MD, R Currò MD, S Facchini, A Tucci MD, Prof M M Reilly MD, Prof H Houlden MD, A Cortese MD), **Department of Neurogedengerative Disease** (Prof J Hardy MD), **and Department of Clinical and Movement Neurosciences** (F Magrinelli MD), **UCL Queen Square Institute of Neurology and The National Hospital for Neurology and Neurosurgery, London, UK; Department of Neurology and Neurosurgery, Montreal Neurological Hospital and Institute, McGill University, Montreal, QC, Canada** (D Pellerin, Prof B Brais MD); **Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy** (R Currò, A Cortese); **William Harvey Research Institute, Queen Mary University of London, London, UK** (A Tucci); **Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA** (Prof D C Bragg PhD, N Sharma PhD); **Department of Human Genetics and Hussman Institute for Human Genomics, University of Miami Miller School of Medicine, Miami, FL, USA** (M C Danzi PhD, Prof S Zuchner MD); **Institute of Medical Genomics, International University of Health and Welfare, Chiba, Japan** (Prof S Tsuji)

Correspondence to:
Dr Andrea Cortese, Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology, University College London, London, WC1N 3BG, UK
andrea.cortese@ucl.ac.uk

Introduction

Repeat expansion diseases are a heterogeneous group of genetic neurological conditions characterised by the presence of short tandem nucleotide repeats in the DNA. These diseases are estimated to affect approximately one in 3000 individuals,[1] although their prevalence might be substantially higher.[2] Notably, the CNS and neuromuscular system are particularly susceptible to the detrimental effects of DNA repeat expansions, as exemplified by the signs and symptoms caused by polyglutamine repeat expansion diseases. These disorders, including Kennedy disease, Huntington's disease, and the most common subtypes of spinocerebellar ataxia, have been well known to general neurologists for more than three decades.

In the past 6 years, technological advancements have unveiled novel pathogenic repeat expansions located within non-coding DNA regions. The presence of these DNA expansions is associated with neurological syndromes, often presenting as epilepsy, cognitive dysfunction, myopathy, neuropathy, ataxia, and movement disorders. Although the conditions stemming from expansions of these novel non-coding DNA repeats are increasingly encountered in clinical practice, they are still unknown to many general neurologists and clinical geneticists.

The primary objective of this Review is to provide a comprehensive overview of the clinical, epidemiological, and molecular features of the adult-onset neurological phenotypes associated with non-coding DNA repeat expansions. By doing so, we aim to provide clinicians with essential knowledge on the accurate diagnosis and management of patients with these diseases, and on genetic counselling for patients and their families.

The diseases and their epidemiology

More than half of the human genome consists of non-coding repetitive elements, including tandem repeats and transposable elements (panel 1).[3] Tandem repeats have the highest mutational rate in the genome, a feature that is evolutionary advantageous because it can favour genetic diversity and facilitate adaptation to changing environments,[4] but has also been implicated in several genetic diseases.[5] Expansions of DNA tandem repeats are known to cause more than 60 monogenic disorders, most of which are primarily neurological.[2]

In the past 6 years, the advent of long-read DNA sequencing and advances in bioinformatics have ushered in a new era for the identification of repeat expansions linked to various neurological conditions. These diseases include familial adult myoclonic epilepsy; neuronal intranuclear inclusion disease; oculopharyngodistal myopathy; spinocerebellar ataxia type 27B; cerebellar ataxia, neuropathy, and vestibular areflexia syndrome (CANVAS); and X-linked dystonia parkinsonism. Most of these repeats are located in non-coding DNA regions, a circumstance that hindered their discovery due to the limitations of sequencing technologies. Carriers of non-coding repeat expansions have been described in many geographical regions, and across various ethnicities, or in specific populations (figure 1).

In the following sections, we describe these diseases according to the mode of inheritance (ie, dominant, recessive, or X-linked) and in chronological order, according to the year of the discovery of the underlying genetic defect (figure 2). Subsequently, we will discuss their differential diagnosis and management. Of note, GCA repeat expansions in the 5′ untranslated region of the *GLS* gene were identified in homozygous or compound heterozygous state with a second missense or nonsense mutation in young patients with glutaminase deficiency, an inborn error of metabolism leading to developmental delay and early-onset progressive ataxia. Also, GGC expansions in the 5′ untranslated region of the *XYLT1* gene were identified in children affected by Baratela-Scott syndrome, a rare disorder characterised by early-onset short stature, facial dysmorphisms, developmental delay, and skeletal dysplasia.[6] However, given our focus on adult-onset neurological diseases, these conditions are not discussed in this Review. CGG expansion in the *ZFHX3* gene was identified in

Panel 1: Glossary of terms

Polyglutamine repeat expansion diseases
A group of genetic disorders characterised by the presence of expansions of a CAG nucleotide triplet, which codes for the amino acid glutamine, within the coding region of the affected gene

Coding DNA
Regions of the genome that contain the instructions for synthesising proteins and that comprise exons, the segments of genes that code for specific protein sequences

Non-coding DNA
Regions of the genome that do not code for proteins, comprising introns within genes, intergenic regions between genes, promoters, 5′ and 3′ untranslated regions, and repetitive elements, including tandem repeats and transposable elements

Tandem repeat
Sequences of DNA in which a short motif, typically consisting of one to six nucleotides, is repeated consecutively multiple times

Transposable element
DNA sequences that can move or transpose within the genome by replicating themselves and reinserting into different locations; the most common type of transposable elements in the human genome are retrotransposons

Retrotransposons
Transposable elements that get inserted in the genome through a copy-and-paste mechanism; by creating a RNA copy of themselves, which is then converted back into DNA, these elements get inserted into a new location in the genome

***Alu* elements**
Short retrotransposons (approximately 300 nucleotides) that constitute approximately 10% of the human genome; they play important roles in regulation of gene expression, genome stability (by facilitating rearrangements and recombination events), and genomic diversity; *Alu* elements have also been implicated in genetic diseases when they get inserted into or disrupt essential genes; *Alu* elements might themselves contain short tandem repeats that can also contribute to human disease when pathologically expanded

Germline instability
The length of DNA repeat sequences changes in the germline cells (sperm or egg cells) of an individual; this variability occurs during the formation of these reproductive cells and can be transmitted to the offspring, leading to an increase in the number of repeats from one generation to the next, which is known as anticipation

Somatic instability
The length of DNA repeat sequences varies within the tissues of an individual over time, resulting in the expansion or contraction of the number of repeat sequences in somatic (non-reproductive) cells, and possibly contributing to the selective involvement of specific regions and neuronal populations

2024 during the revision of this Review.[7,8] However, we do not discuss this disease because the expansion is located in coding DNA.

Genetic and clinical features

Diseases with autosomal dominant inheritance

Familial adult myoclonic epilepsy

Familial adult myoclonic epilepsy, also named benign adult familial myoclonus epilepsy or familial cortical myoclonic tremor with epilepsy, is a highly penetrant autosomal dominant disease with an estimated prevalence of less than 1 case per 35 000 people in Japan and a possible founder effect.[9] Familial adult myoclonic epilepsy typically manifests in adulthood, although onset as early as 11 years has been reported.[10] The disease is clinically characterised by distal myoclonus (cortical tremor), which resembles essential tremor and, although rare, generalised-onset seizures. Intractable seizures and mild cognitive dysfunction have been reported in few people affected by familial adult myoclonic epilepsy.[11] The cortical origin of myoclonus is confirmed by the presence of giant somatosensory evoked potentials, enhanced long-latency electromyography reflexes, and back-averaged EEG time-locked to electromyography. Complex networks engaging sensorimotor cortical and subcortical structures seem to be involved in the pathophysiology of familial adult myoclonic epilepsy.[12]

An intronic pentameric TTTCA repeat expansion in intron 4 of the *SAMD12* gene was identified by long-read DNA sequencing as the cause of familial adult myoclonic epilepsy type 1 in 49 Japanese families years after the locus was initially mapped.[13] The expansion occurs in the polyadenine tail of an *AluSq2* retroelement and can have two different configurations, TTTTA-TTTCA or TTTTA-TTTCA-TTTTA, that range in size from 14 to 3680 repeat units.[13,14] TTTTA expansions are present in approximately 6% of healthy controls of east Asian ancestry, which is not the case for TTTCA expansions, suggesting that the TTTCA motif drives the pathogenic process in familial adult myoclonic epilepsy.[13] Additional rare configurations, including TTTTA-TTTGA-TTTCA[15] and TTTTA-TTTCA-TTTTA-TTTCA,[16] can also lead to familial adult myoclonic epilepsy. Repeat expansions in *SAMD12* have been shown to cause familial adult myoclonic epilepsy type 1 in patients of Canadian or European, Chinese, Indian, Sri Lankan, and Thai descent. All these individuals share the same ancestral haplotype.[16,17]

Following the initial discovery of TTTCA repeat expansions causing familial adult myoclonic epilepsy, expansions of TTTTA-TTTCA or TTTCA in different genes were identified in other disease subtypes, including familial adult myoclonic epilepsy type 2 (*STARD7*)[18] and familial adult myoclonic epilepsy type 3 (*MARCHF6*)[19] in Europeans, familial adult myoclonic epilepsy type 4 (*YEATS2*)[20] in a Thai family, familial adult myoclonic epilepsy type 6 (*TNRC6A*)[13,21] and familial adult myoclonic epilepsy type 7 (*RAPGEF2*)[13,21] in Japanese families, and familial adult myoclonic epilepsy type 8 (*RAI1*)[22] in a Malian family, showing a broad genetic heterogeneity in terms of repeat locus. In most cases, the TTTCA repeat is located in the mid or terminal polyadenine stretch of *Alu* elements.[13] In these diseases, the size of the TTTCA repeat expansion is inversely associated with age of onset.[19] Moreover,

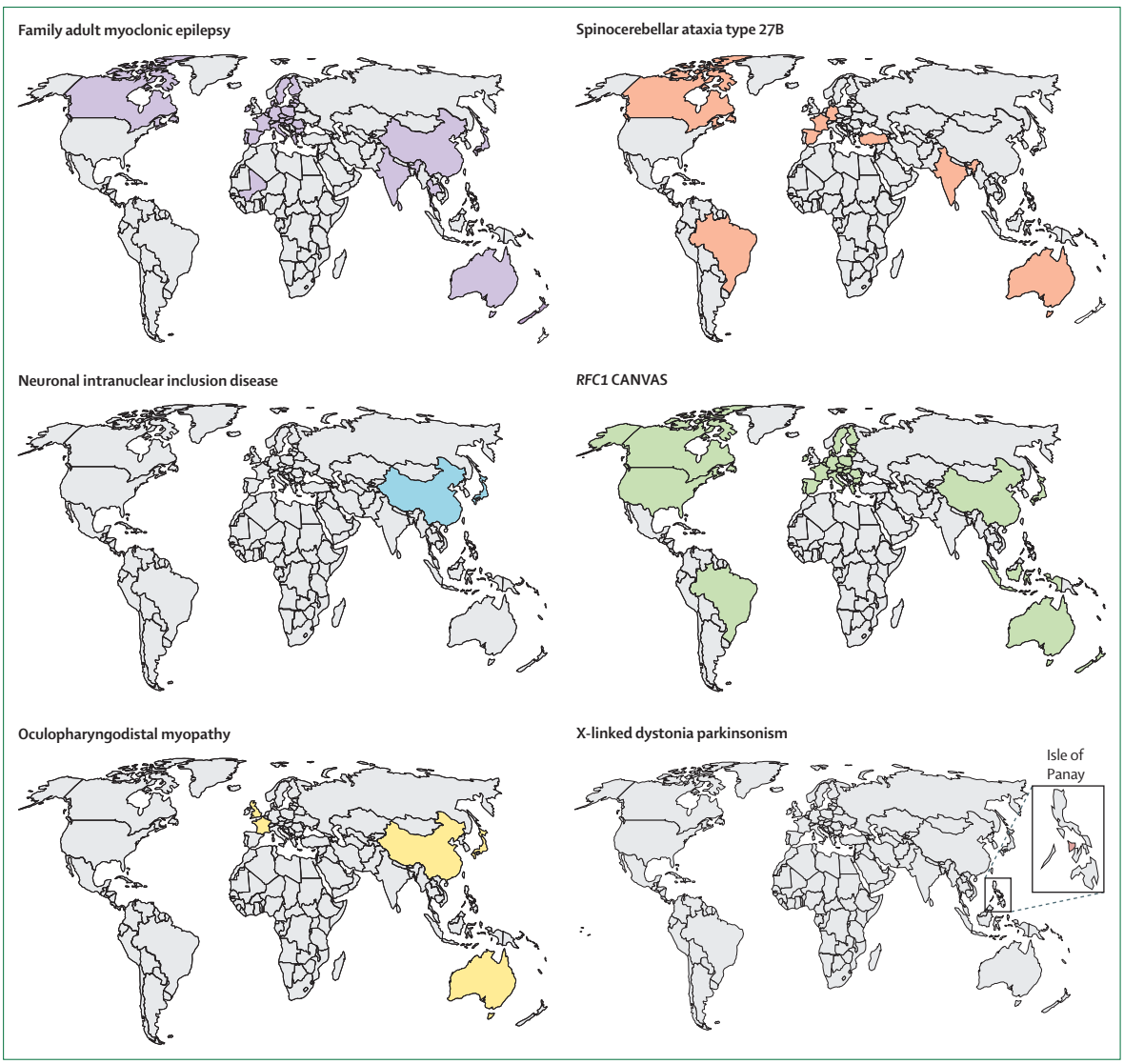

Figure 1: **The geographical regions in which patients with neurological disorders caused by non-coding repeat expansions have been described**

both germline and somatic instabillity have been described.[13]

Loss of *SAMD12* function and the accumulation of toxic RNA foci have been suggested to drive pathogenesis in familial adult myoclonic epilepsy type 1.[13] Notably, RNAs containing UUUCA repeat insertions were previously shown to be toxic in spinocerebellar ataxia type 37, a distinct clinical condition that is also caused by 31–75 intronic TTTCA repeats in the 5′ untranslated region intron 3 of *DAB1*.[23] The identification of the same TTTCA repeat in several ubiquitously expressed genes with distinct functions, from signal transduction (*SAMD12* and *RAPGEF2*), ubiquitination (*MARCHF6*), histone acetylation (*YEATS2*), RNA interference and microRNA-induced gene silencing (*TNRC6A*), to the regulation of circadian clock (*RAI1*), suggests a shared, although still unknown, repeat-dependent and tissue-dependent pathogenic mechanism at least partly unrelated with the specific function of the repeat-containing genes.

Neuronal intranuclear inclusion disease

Neuronal intranuclear inclusion disease is a neurodegenerative disease that is pathologically characterised by the presence of intranuclear ubiquitin and p62 positive inclusions in neurons and astroglial cells. In the past, diagnosis relied on the identification of neuronal intranuclear inclusions post-mortem. Subsequently, the detection of intranuclear eosinophilic inclusions in peripheral tissues, including the skin, has enabled a histological diagnosis of neuronal intranuclear inclusion disease in patients and has led to increased case ascertainment.[24]

The onset of the disease spans from infancy to late adulthood in both familial and sporadic cases, mainly of

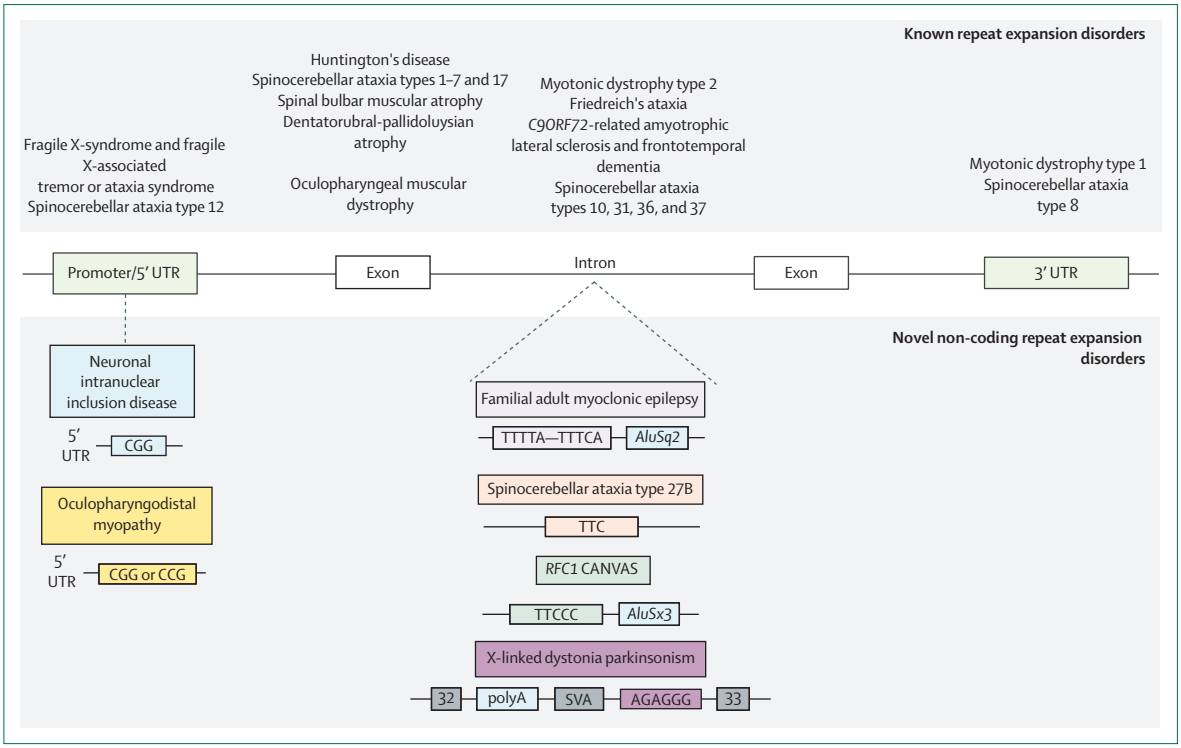

Figure 2: **Genomic location of repeat expansions that cause neurological diseases**
The expanded short tandem repeats are shown with the corresponding location with matching colours: X-linked dystonia parkinsonism (dark purple), familial adult myoclonic epilepsy (light purple), *RFC1* cerebellar ataxia, neuropathy, and vestibular areflexia syndrome (dark green), neuronal intranuclear inclusion disease (light blue), oculopharyngodistal myopathy (yellow), spinocerebellar ataxia type 27B (orange). Representative examples of known repeat expansion disorders are shown in grey above the corresponding location. PolyA=polyadenine. SVA=short interspersed nuclear element, variable number of tandem repeat, and *Alu*. UTR=untranslated region.

Japanese and Chinese ancestry. Although the clinical spectrum of neuronal intranuclear inclusion disease is wide, the syndrome is often dominated by the slow progression of cognitive decline (eg, impaired executive function, abnormal behaviour, and disinhibition) with transient encephalitic-like episodes in patients of east Asian ancestry. Cerebellar ataxia, pyramidal involvement, muscle weakness, sensory impairment, both rest and postural tremor, parkinsonism, dysautonomia, retinopathy, and rare generalised-onset seizures are also reported in some cases.[24,25] Brain MRI typically shows hyperintense signal of the corticomedullary junction on diffusion-weighted imaging. Symmetric white matter T2-weighted and fluid attenuated inversion recovery hyperintensities in the frontal lobes, middle cerebellar peduncles, and in the paravermal area are also commonly observed.[25]

CGG repeat expansion in the 5′ untranslated region of *NOTCH2NLC* was identified as the cause of neuronal intranuclear inclusion disease in more than 70 sporadic and familial cases in Japan.[26–28] This finding was further confirmed in a five-generation Han Chinese family. Pathogenic expansions range from 66 to 525 repeats.[26,27,29] Sequence interruptions can act as disease modifiers.[27] Other neurological presentations of CGG repeat expansion in the 5′ untranslated region of *NOTCH2NLC* expansions include essential tremor,[30] Alzheimer's disease,[26,31] frontotemporal dementia,[31] Parkinson's disease,[26,32] adult-onset leukoencephalopathy,[33] multiple system atrophy,[34] amyotrophic lateral sclerosis,[35] oculopharyngodistal myopathy,[36,37] and, more recently, Charcot-Marie-Tooth disease.[38] Notably, *NOTCH2NLC* expansions are exceedingly rare or absent in individuals of European descent.[37]

In neuronal intranuclear inclusion disease, both toxic RNA foci and polyglycine peptides, which form toxic intranuclear aggregates, have been observed postmortem.[39,40] The gain of function mechanism of RNA and repeat peptides parallels previous observations in fragile X-associated tremor ataxia syndrome, a disorder also caused by 55–200 CGG repeat expansions in the 5′ untranslated region of *FMR1*.[41] There appears to be a pathogenic expansion range, rather than a simple threshold, associated with CGG or CCG expansions, which is different from most other repeat expansion diseases, for which a linear relationship between repeat size and age of onset is usually observed. Large CGG expansions typically lead to gene silencing through DNA methylation and chromatin remodelling, which is detrimental in hemizygous carriers of expansions in *FMR1* (fragile X syndrome) but might be tolerated in autosomal genes associated with neuronal intranuclear inclusion disease or

oculopharyngodistal myopathy, preventing the expression of repeat RNA and polypeptides.[26–29]

Oculopharyngodistal myopathy

Oculopharyngodistal myopathy was first described in 1977 in four families with an autosomal dominant pedigree. It is a rare, adult-onset disease, characterised by progressive ptosis, external ophthalmoplegia, facial weakness, swallowing difficulties, and distal predominant limb weakness. Although most cases have been reported in Japan and China, a few families and sporadic cases have been described also in other regions, including Türkiye and Europe.[42] Muscle biopsy typically reveals chronic myopathic changes, including rimmed vacuoles and intranuclear filamentous inclusions, which are also evident in skin biopsy.

To date, heterozygous CGG or CCG repeat expansions in the 5′ untranslated region of four different genes have been identified in Japanese and Chinese patients with oculopharyngodistal myopathy, including CGG expansions in *LRP12* (oculopharyngodistal myopathy type 1),[26] *GIPC1* (type 2),[43,44] *NOTCH2NLC* (type 3),[36,37] and CCG or GGC (from antisense transcription) expansion in *RILPL1* (type 4).[45,46] Leukodystrophy, peripheral neuropathy, and other neurological manifestations, have been reported in patients with oculopharyngodistal myopathy type 3.[37] Also, CGG or CCG repeat expansions in a bicistronic region containing *LOC642361* (also known as *NUTM2B-AS1*)[26] were identified in a Japanese family with oculopharyngeal myopathy and leukoencephalopathy, thus supporting the existence of a broad phenotypic spectrum of CGG-related disease. More recently, a heterozygous CCG repeat expansion has been identified in the 5′ untranslated region of the *ABCD3* gene (type 5) among Europeans.[47] Similarly to the molecular underpinnings of neuronal intranuclear inclusion disease, in oculopharyngodistal myopathy there seems to be also an interval of pathogenic expansion (for instance, between 85–289 CGG repeats in *LRP12*) while both smaller and larger expansions are tolerated.

Notably, as also described in familial adult myoclonic epilepsy, the identification of CGG or CCG repeats underlying oculopharyngodistal myopathy in several ubiquitously expressed genes involved in diverse cellular processes, including signalling (*LRP12* and *NOTCH2NLC*), scaffolding in the cytoskeleton (*GIPC1*), protein transport and regulation of cell shape and polarity (*RILPL1*), and peroxisome biogenesis (*ABCD3*), suggests that the pathogenic mechanism could be at least partly independent of the repeat-containing genes but might be caused by repeat-dependent toxicity in susceptible muscle tissue. Although the exact mechanism underlying oculopharyngodistal myopathy is largely unknown, RNA-mediated toxicity and protein toxicity (polyglycine peptides) due to repeat-associated non-ATG dependent translation (namely, a form of non-canonical translation initiated at an expanded repeat RNA in the absence of a ATG start codon) have been hypothesised to play a role in myodegeneration.[37,39,43,45,46]

Spinocerebellar ataxia type 27B

The spinocerebellar ataxias are a heterogeneous group of autosomal dominantly inherited disorders characterised by isolated or complex progressive degeneration of the cerebellum, with pyramidal, extrapyramidal, cognitive, peripheral nerve, or retinal involvement. Spinocerebellar ataxias are mainly caused by the expansion of CAG repeats located in the coding regions of multiple genes, which lead to the incorporation of long and aggregate-prone polyglutamine stretches in the open reading frame of these genes. Until recently, a large proportion of patients with isolated cerebellar ataxia were undiagnosed.

TTC repeat expansions (AAG in genomic coordinates) in intron 1 of *FGF14* were identified as a cause of mostly isolated late-onset cerebellar ataxia in 2023. Since point mutations in *FGF14* were already known to cause a form of spinocerebellar ataxia, namely type 27A, the novel disease entity associated with repeat expansions in the same gene was termed type 27B.[48,49]

FGF14 expansions were shown to account for 10–61% of unsolved cases of late-onset ataxia in ethnically diverse cohorts.[48–52] Repeat expansions of at least 250 TTC repeats are deemed pathogenic, although 250–300 TTC expansions appear to be incompletely penetrant.[48,49] Notably, expansions of non-pure TTC motifs, including TTCTCC (AAGAGG in genomic coordinates), appear to be nonpathogenic,[48,52] although their size might be similar or larger than that of pathogenic uninterrupted TTC expansions. Almost a third of cases with spinocerebellar ataxia type 27B do not have a family history, reflecting the high degree of intergenerational instability of the *FGF14* repeat locus.[53]

Patients typically present with a slowly progressive pancerebellar syndrome that is frequently associated with cerebellar oculomotor signs.[54] The disease usually begins between the age of 50 and 70 years. The age of onset inversely correlates, although weakly, with the size of the repeat expansion. Nearly half of the patients experience episodic symptoms at disease onset, which can include diplopia, vertigo, dysarthria, and ataxia. Alcohol intake and exercise are commonly reported triggers. Downbeat nystagmus is observed in about 42% of patients, and visual disturbances, such as oscillopsia, diplopia, and visual blurring, are reported by about 48% of patients.[48] Additional features can include postural tremor, vestibular hypofunction, pyramidal signs, and autonomic dysfunction. Some patients display a mild axonal peripheral sensory or sensorimotor neuropathy.[51,55] Brain MRI can show mild to moderate cerebellar atrophy, which is most pronounced in the vermis. Neuropathological examinations have confirmed the predominant vermian atrophy and loss of cerebellar Purkinje and granule cells and gliosis of the molecular layer.

The intronic repeat expansion in spinocerebellar ataxia type 27B, which is a similar DNA repeat as in Friedreich's ataxia (TTC in *FGF14* for spinocerebellar ataxia type 27B and AAG in *FXN* for Friedreich's ataxia), is thought to cause loss-of-function by interfering with *FGF14* transcription. Preliminary studies in patient-derived post-mortem cerebellum and induced pluripotent stem cell-derived motor neurons have shown a reduction of *FGF14* RNA and protein expression in patients compared with healthy controls.[48]

Autosomal recessive inheritance

CANVAS

The first clinical description of CANVAS dates back to the 1990s, but its genetic cause remained unknown for the next 30 years.[56,57] Patients typically present in their 50s with imbalance, which worsens in the absence of visual guidance. Sensory symptoms and signs appear before the onset of overt ataxia. Muscle bulk, tone, and power are typically preserved. Knee and upper limb reflexes are most often normal or brisk, and ankle reflexes are frequently reduced or absent.[58–60] Approximately a third of patients report head-movement induced oscillopsia,[58,60] while other patients have subclinical involvement of the vestibular system.[58] Cerebellar oculomotor signs, such as gaze-evoked nystagmus, saccadic pursuit, and dysmetric saccades, were also identified in more than half of patients years before subjective complaints of dysarthria and dysphagia. A spasmodic dry cough is fairly typical and can precede the onset of a neurological presentation by up to three decades.[58,60] Autonomic dysfunction is observed in up to a third of patients, although seldom disabling.[58,60] Motor neuron involvement,[61] parkinsonism,[62,63] and cognitive impairment[64] have been described in few patients.

Nerve conduction studies show widespread reduction or absence of sensory nerve action potentials in all patients, but motor conduction studies are typically normal. Brain MRI can show cerebellar atrophy, that is predominant in the vermis at advanced stages in the disease course, and vestibular testing often reveals bilateral vestibular impairment, however both investigations can be initially normal. A few neuropathological studies have shown moderate loss of Purkinje and granule cells, mainly affecting the vermis, with a diffuse neuronal loss in the dorsal roots, trigeminal, and vestibular ganglia.[65,66]

Linkage analysis and whole-genome sequencing in multiple families identified biallelic pentanucleotide TTCCC (AAGGG in genomic coordinates) repeat expansions in intron 2 of *RFC1* as the cause of CANVAS and a frequent cause of late-onset ataxia. The TTCCC expansion maps to the polyadenine tail of an *AluSx3* retroelement and differs in terms of both size and nucleotide sequence from the non-expanded microsatellite containing 11 TTTTC pentanucleotide repeats. The pathogenic expansion usually ranges between 250 and 2000 TTCCC repeats.[56]

Rare pathogenic expansions of different repeat motifs have been described in specific populations, including TTTCC-TTCCC (or AAAGG-AAGGG) in the New Zealand and Cook Island Māori populations[67] and TGTCC (or ACAGG) in Asian-Pacific and Japanese patients.[68] Additional pathogenic expansions include large TTTCC (AAAGG) and TGCCC (AGGGC) repeats, and expansions consisting of two repeat motifs TTGCC-TTCCC (AAGGC-AAGGG) and TTTCC-TTCCC (AAAGG-AAGGG) repeats, which have been identified in Europeans, thus indicating that the size and guanine and cytosine content of repeats might be more important than the exact repeat motif.[69] Notably, the allele frequency of the common TTCCC expansions nears 4% of individuals, suggesting that *RFC1* expansion might be one of the most common recessively inherited neurodegenerative conditions.[56] The possibility of a high disease prevalence is supported by the identification of a biallelic *RFC1* expansion in more than 30% of patients with chronic sensory axonal neuropathy,[70,71] a common condition in people older than 60 years that is generally considered idiopathic.[72]

To date, the mechanism underlying neurodegeneration in people with *RFC1* expansions remains elusive. No changes in *RFC1* mRNA or protein expression were observed in patients' lymphoblasts or fibroblasts, muscle, or post-mortem brain tissue. Moreover, no RNA foci were identified in cerebellum from one autopsy.[56] Although complete loss of *RFC1* appears incompatible with life, the 2023 identification of patients carrying a null variant in *RFC1* on the other allele with respect to TTCCC expansions, leading to decreased mRNA and protein expression, supports an underlying loss-of-function mechanism.[73–75]

X-linked inheritance

X-linked dystonia parkinsonism

X-linked dystonia parkinsonism (also known as Lubag syndrome and formerly dystonia type 3) is a neurodegenerative disorder, which is exclusively observed in individuals whose ancestry can be traced to Panay Island, in the Philippines, due to a presumed founder effect. In men with X-linked dystonia parkinsonism, the average age of onset is 40 years and the average age of death is 56 years.[76] The disease begins most often with focal dystonia, commonly in the legs, and progresses to generalised dystonia, associated with parkinsonism.[76] However, there is evidence that some patients with X-linked dystonia parkinsonism had isolated resting or postural tremor or parkinsonism at onset.[76] Additional features include dysarthria, dysphagia, and cognitive impairment. Women who are heterozygous carriers usually do not develop the full syndrome, although some patients have non-progressive focal dystonia with parkinsonism, albeit milder than in men.[77]

A pathogenic intronic 2·6 kb fragment, comprising a short interspersed nuclear element, a variable number of tandem repeats, and an *Alu* element, is inserted in intron 32 of *TAF1* in individuals with X-linked dystonia parkinsonism.[78–80] The variable number of tandem repeats is a hexameric AGAGGG repeat expansion of 35–52 repeats, the length of which is polymorphic among patients with X-linked dystonia parkinsonism and inversely correlates with age of onset.[78–81] The AGAGGG expansion appears to be the primary cause of the disease, although the participation of other elements of the intronic fragment cannot be excluded.

To date, few studies of postmortem brain tissue from patients with X-linked dystonia parkinsonism have been done. These studies show atrophy of the neostriatum due to a loss of medium spiny neurons.[82] Basal ganglia atrophy, starting from the anterior and medial putamen, and iron accumulation might precede the clinical onset of X-linked dystonia parkinsonism.[83] Furthermore, neuroimaging studies have documented volume loss and functional abnormalities across multiple brain regions,[84] which are consistent with the widespread expression of *TAF1* throughout the CNS.[85] *TAF1* encodes a component of the transcription factor II D complex that mediates transcription by RNA polymerase II.[86] In transcriptomic analyses done in human cells in vitro, three defects in *TAF1* expression have been described: aberrant RNA splicing, increased partial retention of intron 32, and decreased transcription of 3′ exons that reduces the full-length transcript. All of these transcriptional defects were rescued by CRISPR-based excision of the intronic fragment.[80] Therefore, a partial loss-of-function of *TAF1* due to the intronic repeat expansion itself has been hypothesised as the primary mechanism underlying neurodegeneration in X-linked dystonia parkinsonism.[78,80] Nonetheless, additional mechanisms could be at play since missense variants in *TAF1* are known to be associated with intellectual disability,[87] but not parkinsonism or dystonia, and the complete loss of *TAF1* is lethal.

Differential diagnosis

The identification of these novel non-coding repeat expansions has major clinical implication, as patients presenting with common neurological complaints might benefit from genetic testing (panel 2) and these mutations should be incorporated into the clinical reasoning and diagnostic investigation (appendix pp 1–5).

Myoclonus and epilepsy

Myoclonus is a hyperkinetic movement disorder that presents with sudden, brief, involuntary muscle jerks. The initial diagnostic approach is usually guided by the underlying physiology. Myoclonus can be generated in the cortex, subcortex, spinal cord, or peripheral nerves.[93]

After the exclusion of reversible and secondary causes of myoclonus (eg, liver and renal failure and electrolyte and acid-alkaline disturbances), familial adult myoclonic epilepsy should be suspected in adult patients presenting with distal action-induced and

Panel 2: Genetic testing of repeat expansions: methods and limitations

PCR based approaches including repeat-primed PCR and sizing PCR

- Detection of expansions with a known repeat motif
- Cost-effective and available in many diagnostic labs
- Targeted tests (a single repeat locus and repeat expansion can be tested at a time)
- PCR cannot fully amplify expansions, which are large, have a high guanine and cytosine content (eg, *RFC1*), or both, and alternative methods, including Southern blotting, are required for their sizing

Whole-exome sequencing

- Unable to detect most non-coding repeat expansions because of their location (introns) or high GC content (eg, CGG expansion in 5′untranslated region)

Short-read whole-genome sequencing

- High sensitivity and specificity for exonic CAG repeat expansions[88]
- Allows genome-wide profiling of all short-tandem repeats, along with single nucleotide variants and small structural variants
- The read length (approximately 150 nucleotides) is often shorter than the repeat expansion, so short-read whole-genome sequencing is unable to accurately estimate the exact repeat size and motif of large non-coding repeats
- Variable accuracy of bioinformatic tools in predicting size and sequence content of non-coding repeat expansions[89]

Long-read whole-genome sequencing

- Provides reliable information about repeat size and motifs at genome-wide level,[90,91] along with single nucleotide variants, structural variants, and their phasing
- High cost and low availability in genetic labs
- Targeted enrichment methods (eg, CRISPR-Cas9) are available but have variable sequencing yields and accuracy depending on the size and sequence of a specific expansion[69,90]

Non-sequencing based optical genome mapping

- Accurate assessment of all structural variants, including large repeat expansions (>500 nucleotides)[92] at genome-wide level
- Does not provide information on repeat motifs (eg, it cannot distinguish between some non-pathogenic and pathogenic motifs, TTTTA *vs* TTTCA in familial adult myoclonic epilepsy , or TTTTC *vs* TTCCC in *RFC1* cerebellar ataxia, neuropathy, and vestibular areflexia syndrome)
- Has low accuracy for the detection of repeat expansions with fewer than 500 nucleotides
- Cheaper than long-read sequencing

See Online for appendix

posture-induced myoclonus, predominantly affecting the upper limbs, usually exaggerated by sleep deprivation or photostimulation, often in the presence of a dominant family history of myolonus and epilepsy. Generalised-onset seizures with good response to antiepileptic medication are rare in these patients, but their absence does not rule out the disease.[10] Notably, familial adult myoclonic epilepsy can be differentiated from progressive myoclonic epilepsies thanks to its typically non-progressive disease course and the absence of cognitive decline. However, intractable seizures and mild cognitive dysfunction have been reported in few cases with familial adult myoclonic epilepsy type 2.[11]

Familial adult myoclonic epilepsy should also be distinguished from essential tremor. The cortical tremor in patients with familial adult myoclonic epilepsy is more irregular and jerkier than in patients with essential tremor. Also, although alcohol intake might improve essential tremor, it usually worsens cortical tremor and should be avoided (appendix pp 1–5).[94] Despite the unifying pathogenic TTTCA repeat expansion, a broad genetic heterogeneity underlies familial adult myoclonic epilepsy since expansions in *SAMD12*, *YEATS2*, *TNRC6A*, and *RAPGEF2* are typically found in patients from east Asia (except for one reported family of Canadian and European descent with TTTCA expansion in *SAMD12*),[16] while expansions in *STARD7* and *MARCHF6* have been identified in patients of European descent and, although based on a single family, *RAI1* expansions have been found in people from Africa.

***Panel* 3: A case study of a patient with sensory neuropathy**

A woman in her 40s presented to our clinic with burning dysaesthesia in her hands and feet, followed by numbness extending to her extremities. She had no family history of neurological disease or consanguinity. Clinical examination and nerve conduction studies indicated a length-dependent, axonal, sensory neuropathy. A routine laboratory screening for acquired causes of neuropathy was negative. Simultaneously, she complained of dry eyes, dry throat, and chronic cough. A lip biopsy revealed mild lymphocytic and plasma cell infiltration, while extractable nuclear antigen antibodies were negative. The patient received a diagnosis of Sjögren's syndrome-related inflammatory sensory neuropathy and was given hydroxychloroquine. However, the disease progressed and led to gait impairment. About 10 years later, *RFC1* testing revealed the presence of biallelic pathogenic TTCCC (AAGGG) expansions. Importantly, there was no involvement of the cerebellum or vestibular system. *RFC1* expansions are a common cause of sensory neuropathy with cough. In this case, it is likely that the Sjogren's diagnosis was incorrect or coincidental[96] and had little or no effect on the neuropathy, which is relevant because of the potential unnecessary use of immunosuppressive therapies.

Cognitive decline and encephalitic-like episodes

Neuronal intranuclear inclusion disease is a neurodegenerative condition that is almost exclusively observed in patients of east Asian ancestry. The disease typically presents with cognitive decline due to frontal lobe dysfunction and encephalitic-like episodes. A family history, if present, could further orient the diagnostic investigation but is not necessary for the diagnosis of neuronal intranuclear inclusion disease.

Routine laboratory testing, brain MRI, and CSF examination are recommended to rule out other causes of dementia or consciousness impairment, including acquired and genetic leukoencephalopathies and leukodystrophies (appendix pp 1–5). Brain MRI usually shows a high-intensity signal on diffusion-weighted imaging in the corticomedullary junction and T2-weighted hyperintensity in the middle cerebellar peduncles. Elevated proteins, up to approximately 1 g/mL, can be detected in CSF.[24] The identification of ubiquitin-positive and p62-positive intranuclear inclusions on skin biopsy can add evidence towards confirmation of a suspected diagnosis. The definite diagnosis of neuronal intranuclear inclusion disease relies on the identification of CGG repeat expansions in *NOTCH2NLC*.

Oculopharyngeal and distal limb weakness

Patients with oculopharyngodistal myopathy present with onset of ptosis usually aged 10–20 years, followed by external ophthalmoplegia, facial weakness, bulbar involvement, distal limb weakness, and atrophy. A family history might be present, although many cases are sporadic.[42] Needle electromyography can be done to confirm the myopathic nature of this condition. Muscle biopsy is also recommended, since it usually reveals chronic myopathic changes with rimmed vacuoles. Ubiquitin-positive or p62-positive intranuclear or cytoplasmic inclusions are rarely observed.

Oculopharyngodistal myopathy shares some clinical similarities with oculopharyngeal muscular dystrophy, a condition caused by 8–13 GCN (where N represents any nucleotide) repeat expansions in *PABPN1*, which causes facial weakness, ptosis, and dysphagia. However, compared with oculopharyngeal muscular dystrophy, these patients have an earlier onset for ptosis, ophthalmoplegia is more frequent and more severe,[45] and limb weakness predominates distally. Other key differential diagnoses of oculopharyngodistal myopathy include chronic progressive external ophthalmoplegia, different myopathies with distal predominant weakness, and congenital myasthenic syndromes (appendix pp 1–5). Patients from east Asia should be screened for the presence of CGG or CCG repeat expansions in *LRP12*, *GIPC1*, *NOTCH2NLC*, *RILPL1*, and *LOC642361* (also known as *NUTM2B-AS1*), while in individuals of European descent, testing for CCG repeat expansions in the *ABCD3* gene is recommended.

Late-onset ataxia and sensory neuropathy

Ataxia can be caused by the impairment of the cerebellum and the spinocerebellar pathways, the sensory nerves or the posterior columns, or the vestibular system. Medical history and examination generally help to differentiate its origin. Brain and spine MRI, nerve conduction studies, and, if available, vestibular testing should be conducted in all cases for a precise diagnosis and to streamline investigations and genetic testing. The diagnostic approach to late-onset ataxia needs to account for acquired (ie, vascular, neoplasm, inflammatory, toxic-metabolic), degenerative (eg, multiple system atrophy), and genetic causes.[95]

Genetic testing is indicated if the initial investigations are unrevealing and the presentation is consistent with a slowly progressive condition. Clinical features and pattern of inheritance should guide the genetic investigations. Although the common spinocerebellar ataxia (types 1, 2, 3, 6, and 7), Friedreich's ataxia, spastic paraplegia 7, and fragile X-associated tremor ataxia syndrome are routinely tested for, they account for a small proportion of late-onset ataxia cases.

All patients with late-onset ataxia and clinical or neurophysiologic evidence of sensory neuropathy should be tested for biallelic *RFC1* expansions. The *RFC1* test should also be considered early in the diagnostic investigation of an isolated sensory neuropathy or neuronopathy without overt ataxia to avoid misdiagnosis and, potentially, unnecessary immunosuppressive treatments (panel 3).[70,71] The report of chronic cough and the identification of vestibular areflexia further increase the likelihood of a positive *RFC1* test. Conversely, *RFC1* expansions are rare or absent in cases with isolated cerebellar involvement without sensory neuropathy.[97] *RFC1* disease should be differentiated from late-onset Friedreich's ataxia and mitochondrial diseases, including sensory ataxic neuropathy, dysarthria, and ophtalmoparesis caused by biallelic variants in the *POLG* gene and neuropathy, ataxia, and retinitis pigmentosa caused by pathogenic variants in the *MT-ATP6* gene. Conversely, patients with a slowly progressive pan-cerebellar syndrome should undergo testing for *FGF14* expansions. Early episodic symptoms and a family history can also serve to discriminate spinocerebellar ataxia type 27B from *RFC1*-related disease, although in some families carrying *RFC1* expansion, a pseudo-dominant inheritance can be encountered. Also, spinocerebellar ataxia type 27B should be distinguished from episodic ataxias, especially episodic ataxia type 2, and adult-onset spinocerebellar ataxias presenting with a pure cerebellar phenotype, such as spinocerebellar ataxia type 5, 6, 8, and 45 (appendix pp 1–5). Importantly, both *RFC1* and *FGF14* should be considered in the differential diagnosis of multiple system atrophy.

	Polyglutamine repeat expansions	Novel non-coding repeat expansions
Genomic location	Exonic*	Located in non-coding DNA regions, including CGG or CCG expansion at 5′ untranslated region and intronic trinucleotide, pentanucleotide, or hexanucleotide repeat expansions, flanking or inside transposable *Alu* elements
Pathogenic range	Depending on subtype, >30–50 repeat expansions are typically fully pathogenic	Often large (>100 or >1000 repeats), except for some diseases (eg, X-linked dystonia parkinsonism)
Sequence	CAG	Pathogenic repeat usually differs in terms of both size and sequence from the reference satellite (eg, TTTCA in familial adult myoclonic epilepsy and spinocerebellar ataxia type 37,[23] TTCCC in *RFC1* CANVAS)
Penetrance and expressivity	Well explained by premutation and full mutation range, which might be modulated by repeat interruptions; linear correlation between repeat expansion size, age of onset (inverse), and disease severity	Depend on the presence of a mutated repeated unit, which has undergone a change in its nucleotide sequence spanning through all or part of the repeat expansion; correlation between the size of the mutant repeat insertion (rather than total expansion size) and disease severity (eg, TTTCA size in familial adult myoclonic epilepsy)
Family history	Often present; autosomal dominant families with genetic anticipation and parent-of-origin effect	Often absent; recessive inheritance (*RFC1* CANVAS) or dominant inheritance with highly variable penetrance and expressivity in families
Population distribution	Either widely distributed (eg, Huntington's disease) or more frequent in specific populations (eg, spinocerebellar ataxia type 3, dentatorubral-pallidoluysian atrophy)	The repeat expansion is often part of ancestral haplotypes, which might be frequent in specific populations (eg, familial adult myoclonic epilepsy, neuronal intranuclear inclusion disease, X-linked dystonia parkinsonism, oculopharyngodistal myopathy) or shared across different ethnicities (eg, *RFC1* CANVAS, spinocerebellar ataxia type 27B)
Genotype-phenotype correlation	Clinical phenotype depends on both sequence of the expanded repeat and repeat-containing gene; well characterised phenotypes	Characteristic association between the repeat motif and the clinical phenotype (eg, TTTCA in familial adult myoclonic epilepsy, CGG or CCG in oculopharyngodistal myopathy), partly independent from the repeat-containing gene; phenotype spectra are still being characterised
Diagnostic testing	Accurate diagnostic genetic tests are widely available; expansions can be detected from short-read next generation sequencing (whole-genome and whole-exome sequencing)	Diagnostic tests are increasingly available, but mostly still limited to specialised centres; expansions are not detected by use of whole-exome sequencing; short-read whole-genome sequencing can be informative but does not provide accurate sizing and analysis of repeat motif; long-read whole-genome sequencing will probably become the gold standard for diagnostic testing in the future; optical genome mapping can accurately assess all structural variants, including large repeat expansions (>500 nucleotides)[92] at genome-wide level, although it does not provide information on repeat motifs (eg, it cannot distinguish between some non-pathogenic and pathogenic motifs)

CANVAS=cerebellar ataxia, neuropathy, and vestibular areflexia syndrome. *Note that not all the coding (exonic) repeats are polyglutamine stretches (eg, polyalanine stretches in oculopharyngeal muscular dystrophy and polyglycine stretches in spinocerebellar ataxia type 4).

Table 1: **Comparison between neurological disorders caused by polyglutamine repeat expansions and novel non-coding repeat expansions**

	Gene; genomic location (GRCh38 or hg38)	Location in gene	Reference motif	Reference size, number of repeat units	Pathogenic motif*	Pathogenic repeat size, number of repeat units	Pathogenic mechanism	Ethnicity	Main clinical features
Autosomal dominant									
Familial adult myoclonic epilepsy type 1[13,14]	*SAMD12*; chromosome 8 118366813–118366918	Intron	TTTTA (AAAAT)	7–13	TTTCA (TGAAA)	14–3680	RNA-mediated toxicity (RNA foci)	East Asian	Cortical tremor, seizures with generalised motor (tonic-clonic) onset
Familial adult myoclonic epilepsy type 2[18]	*STARD7*; chromosome 2 96197067–96197124	Intron	TTTTA (AAAAT)	12	TTTCA (TGAAA)	150–460	Unknown	European	Cortical tremor, seizures with generalised motor (tonic-clonic) onset
Familial adult myoclonic epilepsy type 3[19]	*MARCHF6*; chromosome 5 10356339–10356411	Intron	TTTTA (TTTTA)	9–20	TTTCA (TTTCA)	668–2814	Somatic genomic rearrangements	European	Cortical tremor, seizures with generalised motor (tonic-clonic) onset
Familial adult myoclonic epilepsy type 4[20]	*YEATS2*; chromsome 3 183712177–183712226	Intron	TTTTA (TTTTA)	7	TTTCA (TTTCA)	962–1262	Unknown	East Asian	Cortical tremor, seizures with generalised motor (tonic-clonic) onset
Familial adult myoclonic epilepsy type 6†[13]	*TNRC6A*; chromosome 16 24613439–24613532	Intron	TTTTA (TTTTA)	18	TTTCA (TTTCA)	27–29	Unknown	East Asian	Cortical tremor, seizures with generalised motor (tonic-clonic) onset
Familial adult myoclonic epilepsy type 7[13]	*RAPGEF2*; chromosome 4 159342527–159342618	Intron	TTTTA (TTTTA)	5–12	TTTCA (TTTCA)	4–19	Unknown	East Asian	Cortical tremor, seizures with generalised motor (tonic-clonic) onset
Familial adult myoclonic epilepsy type 8[22]	*RAI1*; chromosome 17 17808359–17808460	Intron	TTTTA (TTTTA)	16–22	TTTCA (TTTCA)	9–334	Unknown, unchanged *RAI1* expression	African (single large family from Mali)	Cortical tremor, seizures with generalised motor (tonic-clonic) onset
Neuronal intranuclear inclusion disease[26–29]	*NOTCH2NLC*; chromosome 1 149390803–149390842	5′ untranslated region	CGG (CGG)	5–39	CGG (CGG)	66–525	RNA-mediated toxicity (RNA foci), toxic polyglycine peptides	East Asian	Cognitive dysfunction, psychosis, parkinsonism, muscle weakness, sensory disturbances, and pyramidal and cerebellar signs
Oculopharyngodistal myopathy type 1[26]	*LRP12*; chromosome 8 104588973–104588999	5′ untranslated region	CGG (CCG)	9–13	CGG (CCG)	85–289	Unknown	East Asian	Ptosis, external ophthalmoplegia, facial weakness, and pharyngeal and distal limb weakness
Oculopharyngodistal myopathy type 2[43,44]	*GIPC1*; chromosome 19 14496042–14496085	5′ untranslated region	CGG (CCG)	6–31	CGG (CCG)	73–164	Unknown	East Asian	Ptosis, external ophthalmoplegia, facial weakness, and pharyngeal and distal limb weakness
Oculopharyngodistal myopathy type 3[36,37]	*NOTCH2NLC*; chromosome 1 149390803–149390842	5′ untranslated region	CGG (CGG)	6–26	CGG (CGG)	83–674	Unknown	East Asian	Ptosis, external ophthalmoplegia, facial weakness, and pharyngeal and distal limb weakness
Oculopharyngodistal myopathy type 4[45,46]	*RILPL1*; chromosome 12 123533721–123533755	5′ untranslated region and promoter	CCG•CGG (CGG)	12–40	CCG•CGG (CGG)	135–197	RNA-mediated toxicity (RNA foci), toxic polyglycine peptides	East Asian	Ptosis, external ophthalmoplegia, facial weakness, and pharyngeal and distal limb weakness

(Table 2 continues on next page)

	Gene; genomic location (GRCh38 or hg38)	Location in gene	Reference motif*	Reference size, number of repeat units	Pathogenic repeat*	Pathogenic repeat size, number of repeat units	Pathogenic mechanism	Ethnic and geographical distribution	Main clinical features
(Continued from previous page)									
Oculopharyngodistal myopathy type 5[47]	*ABCD3*; chromosome 1 94418389–94518666	5′ untranslated region	CCG (CGG)	7	CCG (CGG)	118–694	Increased expression of repeat containing *ABCD3* transcript	European	Ptosis, external ophthalmoplegia, facial weakness, and pharyngeal and distal limb weakness
Oculopharyngeal myopathy with leukoencephalopathy 1[26]	*LOC642361* (also known as *NUTM2BAS1*); chromosome 10 79826386–79826403	Long non-coding RNA	CGG•CCG (CGG)	6	CGG•CCG (CGG)	Approximately 700	Unknown	East Asian	Oculopharyngodistal myopathy and white matter abnormalities
Spinocerebellar ataxia type 27B[48,49]	*FGF14*; chromosome 13 102161575–102161726	Intron	TTC (AAG)	50	TTC (AAG)	250	Haploinsufficiency	Different ethnicities	Cerebellar ataxia, downbeat nystagmus, episodic symptoms
Autosomal recessive									
RFC1 cerebellar ataxia, neuropathy, and vestibular areflexia syndrome[56]	*RFC1*; chromosome 4 39348425–39348483	Intron	TTTTC (AAAAG)‡	11	TTCCC (AAGGG)‡	250–2000	Unknown, unchanged *RFC1* expression	Different ethnicities	Sensory disturbances, imbalance, oscillopsia, chronic dry cough, dysarthria and dysphagia
X-linked									
X-linked dystonia parkinsonism[79,80]	*TAF1*; chromosome X 71453055–71453129	Intron (retrotransposon)	AGAGGG (AGAGGG)	4	AGAGGG (AGAGGG)	35–52	Altered splicing with intron retention, haploinsufficiency	Filipino	Focal and generalised dystonia, parkinsonism, cognitive dysfunction

For genomic regions that are birectionally transcribed, the two possible nucleotide sequences are separated by a centre dot. *Repeat motifs are indicated as a nucleotide sequence in gene coordinates, according to the transcription sense of the gene, followed by their nucleotide sequence in genomic coordinates, plus strand, in brackets. †FAME5 (1q31·3-q32·2) is not reported in as it is not caused by a repeat expansion but a biallelic single base pair deletion in *CNTN2* gene. ‡Additional rare pathogenic motifs are discussed in the text.

Table 2: **Neurological disorders caused by novel non-coding repeat expansions**

Dystonia and parkinsonism

Dystonia with parkinsonism encompasses a combination of dystonia—a hyperkinetic movement disorder that causes abnormal, often repetitive movements, postures, or both, alongside parkinsonism, which associates bradykinesia with either rest tremor, rigidity, or both. The diagnostic investigation of patients with dystonia and parkinsonism starts with careful phenotyping of the movement disorder. Age, type and tempo of onset, body distribution, temporal pattern of dystonia, presence of other associated features, family history, and levodopa responsiveness are essential information to guide the diagnostic process.

After considering acquired and potentially treatable conditions, including exposure to dopamine receptor blocking medications, and if the preliminary laboratory and imaging findings do not point to a secondary cause, a genetic origin should be considered. The differential diagnoses should include dopamine pathway disorders, inborn errors of metabolism, diseases related to brain metal overload, and recessive and dominant parkinsonisms (appendix pp 1–5), particularly in cases with family history, early onset, or if the clinical phenotype is suggestive of those diseases.[98] Genetic testing for X-linked dystonia parkinsonism should be considered in all men of Filipino ancestry older than 40 years presenting with dystonia, parkinsonism, or a combination of both.

Conclusions and future directions

In the past 6 years, the availability of whole-genome sequencing has fostered the identification of many novel repeat expansions causing neurological disease.[88] Because of their widespread occurrence, neurologists and geneticists must become familiar with the clinical features and molecular causes of these disorders (table 1). These diseases can occur worldwide (eg, *RFC1* CANVAS, spinocerebellar ataxia type 27B), or in specific populations (eg, X-linked dystonia parkinsonism in Filipino men). Importantly, with the exception of familial adult myoclonic epilepsy, which typically manifests in families, patients often do not have a family history,

either because of the recessive mode of inheritance (eg, *RFC1* CANVAS) or highly variable penetrance (up to a third of cases with spinocerebellar ataxia type 27B, neuronal intranuclear inclusion disease, or oculopharyngodistal myopathy are sporadic), so an absence of family history should not preclude further diagnostic investigation.

In the past 6 years, their recognition has provided further evidence of how non-coding repeat expansion diseases, which also include fragile X syndrome (as well as fragile X-associated tremor ataxia syndrome), myotonic dystrophy, *C9orf72* amyotrophic lateral sclerosis and frontotemporal degeneration, and additional subtypes of spinocerebellar ataxia, have molecular and clinical features that differ from those of coding repeat expansion diseases, such as well-known diseases caused by polyglutamine repeat expansions (table 2).[1,6,99,100]

Although specific diagnostic tests are still not widely available, there is intense research to develop sensitive and accurate genetic testing.[90,101] In particular, the gradual adoption in genetic laboratories of long-read sequencing technologies, which provide a more even coverage of these repetitive regions, will lead to increased identification of these mutations, further reducing the diagnostic gap.

Unfortunately, no specific therapies exist for the diseases covered in this Review; hence, management relies on symptomatic treatments. In familial adult myoclonic epilepsy, the treatment is aimed at controlling cortical myoclonus, whereas in neuronal intranuclear inclusion disease, the prevention of concurrent illness is key to avoid encephalitic-like episodes. Both in oculopharyngodistal myopathy and late-onset ataxia, physical and occupational therapy are aimed at preserving functional status, and preventing bulbar or respiratory complications. In patients with *RFC1* CANVAS, tricyclic antidepressants, gabapentinoids, and serotonin and norepinephrine reuptake inhibitors could be considered for neuropathic pain. Pregabalin, amitriptyline, and morphine have shown some beneficial effects in refractory cough, according to anecdotal reports. Also, 4-aminopyridine therapy for downbeat nystagmus and ataxic symptoms in spinocerebellar ataxia type 27B should be considered. In patients with X-linked dystonia parkinsonism, treatment is aimed at improving focal dystonia and parkinsonism (appendix p 6).

Current research efforts are aimed at building international networks to track the natural history of these conditions and identify sensitive biomarkers that could aid in diagnosis, disease monitoring, and the assessment of treatment efficacy.[102] Genome-wide studies of genetic modifiers of disease onset and progression might be also a powerful tool to unravel pathways relevant to their pathogenesis (eg, DNA damage and repair) and to identify therapeutic targets.[103] Although the management of these disorders remains largely symptomatic, we are hopeful that ongoing investigations on the mechanisms by which non-coding repeat expansions cause neurodegeneration will lead to the development of effective therapies, including promising approaches through CRISPR-Cas9 gene editing, small molecule therapies, and antisense oligonucleotides, in the near future.[104]

Search strategy and selection criteria

References included in this Review were identified through searches on PubMed for articles published between Dec 1, 2017, and Feb 29, 2024, and from the references of relevant articles. The main search terms were: "repeat expansion disorder/disease", "non-coding DNA", "Alu element", "microsatellite", "tandem repeat", "next-generation sequencing", "NGS", "whole genome sequencing", "WGS", "long-read", "ataxia", "sensory neuropathy/neuronopathy", "cerebellar ataxia neuropathy and vestibular areflexia syndrome", "CANVAS", "RFC1", "FGF14", "familial adult myoclonic epilepsy", "FAME", "X-linked dystonia parkinsonism", "XDP", "neuronal intranuclear inclusion disease", "NIID", "oculopharyngodistal myopathy", "OPDM", and "SCA27B". There were no language restrictions. The final reference list was generated on the basis of relevance to the topics covered in this Review.

Contributors

EV and AC: conceptualisation and writing the original draft. All authors contributed towards data curation, review, and editing.

Declaration of interests

We declare no competing interests.

Acknowledgments

Our research is funded by the Medical Research Council (MR/T001712/1), Fondazione CARIPLO (1836–2019), the Inherited Neuropathy Consortium, Fondazione Regionale per la Ricerca Biomedica (Regione Lombardia, project ID 1751723), and Progetti di Rilevante Interesse Nazionale (F53D23002330006—20229MMHXP). EV thanks the European Reference Network Research Mobility Fellowship. DP holds a Fellowship award from the Canadian Institutes of Health Research. FM was supported by the MJFF Edmond J Safra Clinical Research Fellowship in Movement Disorders. We thank Chris Record and Valentine Perrain for their help with the clinical vignette.

References

1 Paulson H. Repeat expansion diseases. In: Geschwind DH, Paulson HL, Klein C, eds. Handbook of Clinical Neurology, part 1. Netherlands: Elsevier, 2018: 105–23.

2 Ibañez K, Jadhav B, Facchini S, et al. Population frequency of repeat expansions indicates increased disease prevalence estimates across different populations. *medRxiv* 2023; published online July 6. https://doi.org/10.1101/2023.07.03.23292162 (preprint).

3 Nurk S, Koren S, Rhie A, et al. The complete sequence of a human genome. *Science* 2022; **376:** 44–53.

4 Salem A-H, Ray DA, Xing J, et al. *Alu* elements and hominid phylogenetics. *Proc Natl Acad Sci USA* 2003; **100:** 12787–91.

5 Verbiest M, Maksimov M, Jin Y, Anisimova M, Gymrek M, Bilgin Sonay T. Mutation and selection processes regulating short tandem repeats give rise to genetic and phenotypic diversity across species. *J Evol Biol* 2023; **36:** 321–36.

6 Depienne C, Mandel J-L. 30 years of repeat expansion disorders: what have we learned and what are the remaining challenges? *Am J Hum Genet* 2021; **108:** 764–85.

7 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.

8 Chen Z, Gustavsson EK, Macpherson H, et al. Adaptive long-read sequencing reveals GGC repeat expansion in *ZFHX3* associated with spinocerebellar ataxia type 4. *Mov Disord* 2024; **39:** 486–97.

9 Henden L, Freytag S, Afawi Z, et al. Identity by descent fine mapping of familial adult myoclonus epilepsy (FAME) to 2p11.2–2q11.2. *Hum Genet* 2016; **135:** 1117–25.

10 Ikeda A, Kakigi R, Funai N, Neshige R, Kuroda Y, Shibasaki H. Cortical tremor: a variant of cortical reflex myoclonus. *Neurology* 1990; **40:** 1561–65.

11 Guerrini R, Bonanni P, Patrignani A, et al. Autosomal dominant cortical myoclonus and epilepsy (ADCME) with complex partial and generalized seizures: a newly recognized epilepsy syndrome with linkage to chromosome 2p11.1–q12.2. *Brain* 2001; **124:** 2459–75.

12 Dubbioso R, Striano P, Tomasevic L, et al. Abnormal sensorimotor cortex and thalamo-cortical networks in familial adult myoclonic epilepsy type 2: pathophysiology and diagnostic implications. *Brain Commun* 2022; **4:** fcac037.

13 Ishiura H, Doi K, Mitsui J, et al. Expansions of intronic TTTCA and TTTTA repeats in benign adult familial myoclonic epilepsy. *Nat Genet* 2018; **50:** 581–90.

14 Mizuguchi T, Toyota T, Miyatake S, et al. Complete sequencing of expanded *SAMD12* repeats by long-read sequencing and Cas9-mediated enrichment. *Brain* 2021; **144:** 1103–17.

15 Cen Z, Chen Y, Yang D, et al. Intronic $(TTTGA)_n$ insertion in *SAMD12* also causes familial cortical myoclonic tremor with epilepsy. *Mov Disord* 2019; **34:** 1571–76.

16 Maroilley T, Tsai M, Mascarenhas R, et al. A novel FAME1 repeat configuration in a European family identified using a combined genomics approach. *Epilepsia Open* 2023; **8:** 659–65.

17 Cherian A, Divya KP, Krishnan ARS. Familial adult myoclonus epilepsy: a pragmatic approach. *Acta Neurol Belg* 2023; **124:** 389–94.

18 Corbett MA, Kroes T, Veneziano L, et al. Intronic ATTTC repeat expansions in *STARD7* in familial adult myoclonic epilepsy linked to chromosome 2. *Nat Commun* 2019; **10:** 4920.

19 Florian RT, Kraft F, Leitão E, et al. Unstable TTTTA/TTTCA expansions in *MARCH6* are associated with familial adult myoclonic epilepsy type 3. *Nat Commun* 2019; **10:** 4919.

20 Yeetong P, Pongpanich M, Srichomthong C, et al. TTTCA repeat insertions in an intron of *YEATS2* in benign adult familial myoclonic epilepsy type 4. *Brain* 2019; **142:** 3360–66.

21 Lei XX, Liu Q, Lu Q, et al. TTTCA repeat expansion causes familial cortical myoclonic tremor with epilepsy. *Eur J Neurol* 2019; **26:** 513–18.

22 Yeetong P, Dembélé ME, Pongpanich M, et al. Pentanucleotide repeat insertions in RAI1 cause benign adult familial myoclonic epilepsy type 8. *Mov Disord* 2023; **39:** 164–72.

23 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.

24 Sone J, Mori K, Inagaki T, et al. Clinicopathological features of adult-onset neuronal intranuclear inclusion disease. *Brain* 2016; **139:** 3170–86.

25 Bao L, Zuo D, Li Q, Chen H, Cui G. Current advances in neuronal intranuclear inclusion disease. *Neurol Sci* 2023; **44:** 1881–89.

26 Ishiura H, Shibata S, Yoshimura J, et al. Noncoding CGG repeat expansions in neuronal intranuclear inclusion disease, oculopharyngodistal myopathy and an overlapping disease. *Nat Genet* 2019; **51:** 1222–32.

27 Sone J, Mitsuhashi S, Fujita A, et al. Long-read sequencing identifies GGC repeat expansions in *NOTCH2NLC* associated with neuronal intranuclear inclusion disease. *Nat Genet* 2019; **51:** 1215–21.

28 Tian Y, Wang J-L, Huang W, et al. Expansion of human-specific GGC repeat in neuronal intranuclear inclusion disease-related disorders. *Am J Hum Genet* 2019; **105:** 166–76.

29 Tai H, Wang A, Zhang Y, et al. Clinical features and classification of neuronal intranuclear inclusion disease. *Neurol Genet* 2023; **9:** e200057.

30 Sun QY, Xu Q, Tian Y, et al. Expansion of GGC repeat in the human-specific *NOTCH2NLC* gene is associated with essential tremor. *Brain* 2020; **143:** 222–33.

31 Jiao B, Zhou L, Zhou Y, et al. Identification of expanded repeats in *NOTCH2NLC* in neurodegenerative dementias. *Neurobiol Aging* 2020; **89:** 142.

32 Ma D, Tan YJ, Ng ASL, et al. Association of *NOTCH2NLC* repeat expansions with Parkinson's disease. *JAMA Neurol* 2020; **77:** 1559–63.

33 Liu YH, Chou YT, Chang FP, et al. Neuronal intranuclear inclusion disease in patients with adult-onset non-vascular leukoencephalopathy. *Brain* 2022; **145:** 3010–21.

34 Fang P, Yu Y, Yao S, et al. Repeat expansion scanning of the *NOTCH2NLC* gene in patients with multiple system atrophy. *Ann Clin Transl Neurol* 2020; **7:** 517–26.

35 Yuan Y, Liu Z, Hou X, et al. Identification of GGC repeat expansion in the *NOTCH2NLC* gene in amyotrophic lateral sclerosis. *Neurology* 2020; **95:** e3394–405.

36 Ogasawara M, Iida A, Kumutpongpanich T, et al. CGG expansion in *NOTCH2NLC* is associated with oculopharyngodistal myopathy with neurological manifestations. *Acta Neuropathol Commun* 2020; **8:** 204.

37 Yu J, Deng J, Guo X, et al. The GGC repeat expansion in *NOTCH2NLC* is associated with oculopharyngodistal myopathy type 3. *Brain* 2021; **144:** 1819–32.

38 Liao Y-C, Chang F-P, Huang H-W, et al. GGC repeat expansion of *NOTCH2NLC* in Taiwanese patients with inherited neuropathies. *Neurology* 2021; **98:** e199–206.

39 Deng J, Zhou B, Yu J, et al. Genetic origin of sporadic cases and RNA toxicity in neuronal intranuclear inclusion disease. *J Med Genet* 2021; **59:** 462–69.

40 Boivin M, Deng J, Pfister V, et al. Translation of GGC repeat expansions into a toxic polyglycine protein in NIID defines a novel class of human genetic disorders: the polyG diseases. *Neuron* 2021; **109:** 1825–35.

41 Fu YH, Kuhl DPA, Pizzuti A, et al. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 1991; **67:** 1047–58.

42 Ishiura H, Tsuji S, Toda T. Recent advances in CGG repeat diseases and a proposal of fragile X-associated tremor/ataxia syndrome, neuronal intranuclear inclusion disease, and oculophryngodistal myopathy (FNOP) spectrum disorder. *J Hum Genet* 2023; **68:** 169–74.

43 Deng J, Yu J, Li P, et al. Expansion of GGC repeat in *GIPC1* is associated with oculopharyngodistal myopathy. *Am J Hum Genet* 2020; **106:** 793–804.

44 Xi J, Wang X, Yue D, et al. 5′ UTR CGG repeat expansion in *GIPC1* is associated with oculopharyngodistal myopathy. *Brain* 2021; **144:** 601–14.

45 Yu J, Shan J, Yu M, et al. The CGG repeat expansion in *RILPL1* is associated with oculopharyngodistal myopathy type 4. *Am J Hum Genet* 2022; **109:** 533–41.

46 Zeng Y, Yang K, Du G, et al. GGC repeat expansion of *RILPL1* is associated with oculopharyngodistal myopathy. *Ann Neurol* 2022; **92:** 512–26.

47 Cortese A, Beecroft SJ, Facchini S, et al. A CCG expansion in *ABCD3* causes oculopharyngodistal myopathy in individuals of European ancestry. *medRxiv* 2023; published online Oct 10. https://doi.org/10.1101/2023.10.09.23296582 (preprint).

48 Pellerin D, Danzi MC, Wilke C, et al. Deep intronic *FGF14* GAA repeat expansion in late-onset cerebellar ataxia. *N Engl J Med* 2023; **388:** 128–41.

49 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:** 105–19.

50 Bonnet C, Pellerin D, Roth V, et al. Ca MR. Optimized testing strategy for the diagnosis of GAA-FGF14 ataxia/spinocerebellar ataxia 27B. *Sci Rep* 2023; **13:** 9737.

51 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:** 1950–56.

52 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:** 1557–58.

53 Pellerin D, Del Gobbo G, Couse M, et al. A common flanking variant is associated with enhanced meiotic stability of the *FGF14*-SCA27B locus. *bioRxiv* 2023; published online June 30. https://doi.org/10.1101/2023.05.11.540430 (preprint).

54 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:** 4144–57.

55 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* 2023; **95:** 175–79.

56 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:** 649–58.

57 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:** 151–65.

58 Cortese A, Tozza S, Yau WY, et al. Cerebellar ataxia, neuropathy, vestibular areflexia syndrome due to RFC1 repeat expansion. *Brain* 2020; **143:** 480–90.

59 Montaut S, Diedhiou N, Fahrer P, et al. Biallelic *RFC1*-expansion in a French multicentric sporadic ataxia cohort. *J Neurol* 2021; **268:** 3337–43.

60 Cortese A, Curro' R, Vegezzi E, Yau WY, Houlden H, Reilly MM. Cerebellar ataxia, neuropathy and vestibular areflexia syndrome (CANVAS): genetic and clinical aspects. *Pract Neurol* 2021; **22:** 14–18.

61 Huin V, Coarelli G, Guemy C, et al. Motor neuron pathology in CANVAS due to *RFC1* expansions. *Brain* 2021; **145:** 2121–32.

62 Sullivan R, Yau WY, Chelban V, et al. *RFC1*-related ataxia is a mimic of early multiple system atrophy. *J Neurol Neurosurg Psychiatry* 2021; **92:** 444–46.

63 Ylikotila P, Sipilä J, Alapirtti T, et al. Association of biallelic *RFC1* expansion with early-onset Parkinson's disease. *Eur J Neurol* 2023; **30:** 1256–61.

64 Korpioja A, Krüger J, Hurme-Niiranen A, et al. Cognitive impairment is not uncommon in patients with biallelic *RFC1* AAGGG repeat expansion, but the expansion is rare in patients with cognitive disease. *Parkinsonism Relat Disord* 2022; **103:** 98–101.

65 Szmulewicz DJ, Merchant SN, Halmagyi GM. Cerebellar ataxia with neuropathy and bilateral vestibular areflexia syndrome: a histopathologic case report. *Otol Neurotol* 2011; **32:** e63–65.

66 Szmulewicz DJ, McLean CA, Rodriguez ML, et al. Dorsal root ganglionopathy is responsible for the sensory impairment in CANVAS. *Neurology* 2014; **82:** 1410–15.

67 Beecroft SJ, Cortese A, Sullivan R, et al. A Māori specific *RFC1* pathogenic repeat configuration in CANVAS, likely due to a founder allele. *Brain* 2020; **143:** 2673–80.

68 Scriba CK, Beecroft SJ, Clayton JS, et al. A novel *RFC1* repeat motif (ACAGG) in two Asia-Pacific CANVAS families. *Brain* 2020; **143:** 2904–10.

69 Dominik N, Magri S, Currò R, et al. Normal and pathogenic variation of *RFC1* repeat expansions: implications for clinical diagnosis. *Brain* 2023; **146:** 5060–69.

70 Currò R, Salvalaggio A, Tozza S, et al. *RFC1* expansions are a common cause of idiopathic sensory neuropathy. *Brain* 2021; **144:** 1542–50.

71 Tagliapietra M, Cardellini D, Ferrarini M, et al. *RFC1* AAGGG repeat expansion masquerading as chronic idiopathic axonal polyneuropathy. *J Neurol* 2021; **268:** 4280–90.

72 Smith AG, Singleton JR. The diagnostic yield of a standardized approach to idiopathic sensory-predominant neuropathy. *Arch Intern Med* 2004; **164:** 1021–25.

73 Ronco R, Perini C, Currò R, et al. Truncating variants in *RFC1* in cerebellar ataxia, neuropathy, and vestibular areflexia syndrome. *Neurology* 2023; **100:** e543–54.

74 Benkirane M, Da Cunha D, Marelli C, et al. *RFC1* nonsense and frameshift variants cause CANVAS: clues for an unsolved pathophysiology. *Brain* 2022; **145:** 3770–75.

75 Arteche-López A, Avila-Fernandez A, Damian A, et al. New cerebellar ataxia, neuropathy, vestibular areflexia syndrome cases are caused by the presence of a nonsense variant in compound heterozygosity with the pathogenic repeat expansion in the *RFC1* gene. *Clin Genet* 2023; **103:** 236–41.

76 Lee LV, Rivera C, Teleg RA, et al. The unique phenomenology of sex-linked dystonia parkinsonism (XDP, DYT3, "Lubag"). *Int J Neurosci* 2011; **121** (suppl 1)**:** 3–11.

77 Evidente VGH, Nolte D, Niemann S, et al. Phenotypic and molecular analyses of X-linked dystonia-parkinsonism ("lubag") in women. *Arch Neurol* 2004; **61:** 1956–59.

78 Makino S, Kaji R, Ando S, et al. Reduced neuron-specific expression of the *TAF1* gene is associated with X-linked dystonia-parkinsonism. *Am J Hum Genet* 2007; **80:** 393–406.

79 Bragg DC, Mangkalaphiban K, Vaine CA, et al. Disease onset in X-linked dystonia-parkinsonism correlates with expansion of a hexameric repeat within an SVA retrotransposon in *TAF1*. *Proc Natl Acad Sci USA* 2017; **114:** E11020–28.

80 Aneichyk T, Hendriks WT, Yadav R, et al. Dissecting the causal mechanism of X-linked dystonia-parkinsonism by integrating genome and transcriptome assembly. *Cell* 2018; **172:** 897–909.

81 Westenberger A, Reyes CJ, Saranza G, et al. A hexanucleotide repeat modifies expressivity of X-linked dystonia parkinsonism. *Ann Neurol* 2019; **85:** 812–22.

82 Goto S, Kawarai T, Morigaki R, et al. Defects in the striatal neuropeptide Y system in X-linked dystonia-parkinsonism. *Brain* 2013; **136:** 1555–67.

83 Hanssen H, Diesta CCE, Heldmann M, et al. Basal ganglia atrophy as a marker for prodromal X-linked dystonia-parkinsonism. *Ann Neurol* 2023; **93:** 999–1011.

84 Brüggemann N, Heldmann M, Klein C, et al. Neuroanatomical changes extend beyond striatal atrophy in X-linked dystonia parkinsonism. *Parkinsonism Relat Disord* 2016; **31:** 91–97.

85 Capponi S, Stöffler N, Penney EB, et al. Dissection of *TAF1* neuronal splicing and implications for neurodegeneration in X-linked dystonia-parkinsonism. *Brain Commun* 2021; **3:** fcab253.

86 Schier AC, Taatjes DJ. Structure and mechanism of the RNA polymerase II transcription machinery. *Genes Dev* 2020; **34:** 465–88.

87 O'Rawe JA, Wu Y, Dörfel MJ, et al. *TAF1* variants are associated with dysmorphic features, intellectual disability, and neurological manifestations. *Am J Hum Genet* 2015; **97:** 922–32.

88 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:** 234–45.

89 Dolzhenko E, Bennett MF, Richmond PA, et al. ExpansionHunter Denovo: a computational method for locating known and novel repeat expansions in short-read sequencing data. *Genome Biol* 2020; **21:** 102.

90 Stevanovski I, Chintalaphani SR, Gamaarachchi H, et al. Comprehensive genetic diagnosis of tandem repeat expansion disorders with programmable targeted nanopore sequencing. *Sci Adv* 2022; **8:** eabm5386.

91 Nakamura H, Doi H, Mitsuhashi S, et al. Long-read sequencing identifies the pathogenic nucleotide repeat expansion in *RFC1* in a Japanese case of CANVAS. *J Hum Genet* 2020; **65:** 475–80.

92 Facchini S, Dominik N, Manini A, et al. Optical genome mapping enables detection and accurate sizing of *RFC1* repeat expansions. *Biomolecules* 2023; **13:** 1546.

93 Zutt R, van Egmond ME, Elting JW, et al. A novel diagnostic approach to patients with myoclonus. *Nat Rev Neurol* 2015; **11:** 687–97.

94 Striano P, de Falco FA, Minetti C, Zara F. Familial benign nonprogressive myoclonic epilepsies. *Epilepsia* 2009; **50** (suppl 5)**:** 37–40.

95 Bogdan T, Wirth T, Iosif A, et al. Unravelling the etiology of sporadic late-onset cerebellar ataxia in a cohort of 205 patients: a prospective study. *J Neurol* 2022; **269:** 6354–65.

96 Fernández-Eulate G, Debs R, Maisonobe T, et al. Sjögren syndrome and *RFC1*-CANVAS sensory ganglionopathy: co-occurrence or misdiagnosis? *J Neurol* 2023; **270:** 460–65.

97 Hadjivassiliou M, Currò R, Beauchamp N, et al. Can CANVAS due to *RFC1* biallelic expansions present with pure ataxia? *J Neurol Neurosurg Psychiatry* 2023; **95:** 171–74.

98 Morales-Briceno H, Fung VSC, Bhatia KP, Balint B. Parkinsonism and dystonia: clinical spectrum and diagnostic clues. *J Neurol Sci* 2022; **433:** 120016.

99 Hannan AJ. Tandem repeats mediating genetic plasticity in health and disease. *Nat Rev Genet* 2018; **19:** 286–98.

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

101 Miyatake S, Koshimizu E, Fujita A, et al. Rapid and comprehensive diagnostic method for repeat expansion diseases using nanopore sequencing. *NPJ Genom Med* 2022; **7:** 62.

102 Coarelli G, Coutelier M, Durr A. Autosomal dominant cerebellar ataxias: new genes and progress towards treatments. *Lancet Neurol* 2023; **22:** 735–49.

103 Jones L, Houlden H, Tabrizi SJ. DNA repair in the trinucleotide repeat disorders. *Lancet Neurol* 2017; **16:** 88–96.

104 Wang ET, Freudenreich CH, Gromak N, Jain A, Todd PK, Nagai Y. What repeat expansion disorders can teach us about the central dogma. *Mol Cell* 2023; **83:** 324–29.