

REVIEW

Open Access



The polyG diseases: a new disease entity

Tongling Liufu^{1†}, Yilei Zheng^{2†}, Jiayi Yu¹, Yun Yuan¹, Zhaoxia Wang^{1,3}, Jianwen Deng^{1,3*} and Daojun Hong^{2,4*}

Abstract

Recently, inspired by the similar clinical and pathological features shared with fragile X-associated tremor/ataxia syndrome (FXTAS), abnormal expansion of CGG repeats in the 5' untranslated region has been found in neuronal intranuclear inclusion disease (NIID), oculopharyngeal myopathy with leukoencephalopathy (OPML), and oculopharyngodistal myopathy (OPDMs). Although the upstream open reading frame has not been elucidated in OPML and OPDMs, polyglycine (polyG) translated by expanded CGG repeats is reported to be as a primary pathogenesis in FXTAS and NIID. Collectively, these findings indicate a new disease entity, the polyG diseases. In this review, we state the common clinical manifestations, pathological features, mechanisms, and potential therapies in these diseases, and provide preliminary opinions about future research in polyG diseases.

Keywords: PolyG diseases, Fragile X-associated tremor/ataxia syndrome, Neuronal intranuclear inclusion disease, Oculopharyngeal myopathy with leukoencephalopathy, Oculopharyngodistal myopathy

Introduction

Short tandem repeats (STRs) are nucleotide repeats located in both coding and non-coding regions throughout the human genome [103]. STR variants have been reported as the cause of a series of neurodegenerative diseases, including myotonic dystrophies (DM1 and DM2), fragile X tremor/ataxia syndrome (FXTAS), some spinocerebellar ataxias (SCAs), and chromosome 9 open reading frame 72 (*C9ORF72*)-related amyotrophic lateral sclerosis (ALS) [21]. More and more studies suggest that STR variants are an important and potential genetic cause of neurodegenerative diseases.

Multiple neurodegenerative diseases have been identified with trinucleotide repeat expansions. The first trinucleotide repeat expansion in a non-coding region on chromosome X was found in patients with fragile X syndrome in 1991 through linkage analysis [114], meanwhile

the first trinucleotide repeat expansion in a coding region on androgen receptor gene was identified in Kennedy disease (KD) [56]. With the advances of next-generation sequencing technology, especially in long-read sequencing (LRS), CGG repeat expansion in non-coding regions was found to be associated with a group of previously undetermined late-onset neurodegenerative diseases, including neuronal intranuclear inclusion disease (NIID) [18, 43, 110], oculopharyngeal myopathy with leukoencephalopathy (OPML) [43], and oculopharyngodistal myopathy (OPDM) [19, 43, 129].

The above disorders with non-coding CGG repeat expansion in the 5' untranslated region (5'UTR) share substantial overlap in clinical, neuroimaging, and pathological features, so we hypothesize that they have similar pathogenic mechanisms (Table 1). Polyglycine (polyG) originating from the expanded CGG repeats, which can be translated into a gain-of-function toxic protein, has been reported as an important pathogenic mechanism for FXTAS and NIID [27, 94, 134]. Moreover, preliminary pathological investigations have shown polyG-positive intranuclear inclusions (NIIs) may be deposited in the biopsy specimens from patients with OPDM types 3–4 [127, 129], suggesting a similar pathogenesis with FXTAS and NIID. Accordingly, these disorders have been

[†]Tongling Liufu and Yilei Zheng have contributed equally to this work.

*Correspondence: jianwendeng@pkuhf.com; hongdaojun@hotmail.com

¹ Department of Neurology, Peking University First Hospital, Beijing, China

² Department of Neurology, The First Affiliated Hospital of Nanchang University, Nanchang, China

Full list of author information is available at the end of the article



Table 1 Phenotypic features of polyG diseases

	FXTAS	NIID	OPML	OPDM			
				OPDM1	OPDM2	OPDM3	OPDM4
Mode of inheritance	XL	AD	AD	AD	AD	AD	AD
Affected genes	<i>FMR1</i>	<i>NOTCH2NLC</i>	<i>NUTM2B-AS1</i>	<i>LRP12</i>	<i>GIPC1</i>	<i>NOTCH2NLC</i>	<i>RILPL1</i>
Gene location	Xq27.3	1q21.2	10q22.3	8q22.3	19p13.12	1q21.2	12q24.31
CGG expansion location	5'UTR	5'UTR	Non-coding Transcript	5'UTR	5'UTR	5'UTR	5'UTR
Physiological repeat numbers	5–50	7–40?	3–16	13–45	12–32	7–40?	9–16
Pathological repeat numbers	55–200	60–300	50–60	85–289	73–164	60–300	139–197
Age of onset (years)	> 50	most > 50	15–40	30–50	20–60	20–50	15–40
<i>Clinical features</i>							
Movement disorders	+++	+	+	+	+	+	+
Cognitive deficit	+	+++	±	–	+	+	–
Autonomic dysfunction	+++	+++	+	–	–	+	–
Muscle-weakness	+	+	+	++	++	++	++
Others	neuropsychosis	encephalitic episodes	myocardiopathy	myocardiopathy	–	–	–
<i>Neuroimage features</i>							
Brian atrophy	+	+	+	–	–	+	–
Leukoencephalopathy	+	+	+	–	+	+	–
Ribbon sign	±	+	±	–	–	±	–
<i>Pathology</i>							
Eosinophilic inclusions	+	+++	–	+	+	+	+
Tubulofilamentous inclusion	+	+	–	+	+	+	+
Ubiquitin- or p62- positive	+	+	N/A	+	+	+	+
PolyG-positive	+	+	(polyG)?	(polyG)?	(polyG)?	+	(polyG)?
Rimmed vacuoles in muscle	–	±	–	+	+	+	+

AD Autosomal dominant; XL X-linked dominant; AR Autosomal recessive; N/A not available; UTR untranslated region

+++ , most positive; ++ , some positive; + , a few positive; ± , occasionally positive; – , negative

FXTAS fragile X-associated tremor ataxia syndrome; NIID neuronal intranuclear inclusion disease; OPML oculopharyngeal myopathy with leukoencephalopathy; OPDM1 oculopharyngodistal myopathy type 1–4; polyG polyGlycine peptide toxicity

classified as a novel entity of disease—the polyG diseases [10, 134]. The polyG diseases are a new disease entity, which are characterized by the polyG protein being deposited in the nucleus with a common genetic cause: CGG repeat expansions in the 5'UTR regions.

In this review, we provide a summary of recent findings of polyG diseases, and highlight the similarities and differences among these diseases. The pathogenic mechanism across different polyG diseases, future research directions, potential therapeutic strategies and challenges will be discussed.

Polyglycine(G) disorders

Fragile X-associated tremor/ataxia syndrome (FXTAS)

FXTAS is a late-onset, X-linked, neurodegenerative disease characterized by intention tremor, cerebellar ataxia, and cognitive decline, which principally affects males [9, 32]. This disorder is caused by premutation with CGG repeat expansion (55–200 repeats) in the 5'UTR of the fragile X mental retardation 1 (*FMR1*) gene [35]. *FMR1*

is located on Xq27.3 and encodes the RNA-binding protein *FMR1* protein (FMRP) [98]. Methylation appears in full mutation CGG expansions (>200 repeats), leading to transcriptional silencing with consequent deficit FMRP levels in the cells and resulting in fragile X syndrome (FXS), one of the most common inherited forms of intellectual disability and an autism spectrum disorder [23, 89]. Women who carry the expansion will experience lower than 20% risk in having FXTAS due to random inactivation in one of two X chromosomes [34]. Males with CGG repeat numbers less than 71 have low penetrance of FXTAS [71]. In the general population, the prevalence of premutation in CGG repeat expansions is approximately 1 in 300 females and 1 in 850 males [21].

More than one in three adult male premutation carriers present with the neurologic syndrome, including progressive gait ataxia, tremor, cognitive decline, parkinsonism, psychological disorders, and generalized brain atrophy over 50 years of age [35]. Typical neuroimaging changes in FXTAS include symmetrical T2-weighted and

fluid attenuated inversion recovery (FLAIR) sequence hyperintense changes in the middle cerebellar peduncle (MCP sign), extensive white matter lesions, brain atrophy, and diffusion-weighted imaging (DWI) sequence hyperintensity at the cortico-medullary junction [45].

Premutation carriers used to be regarded as a phenotypic variant of FXS with normal cognitive abilities for many years. However, recent pathological examination suggested that FXTAS is a different disease entity: a type of inclusion disease with eosinophilic NIIs in both neurons and astrocytes throughout the brain [33]. The NIIs usually contain *FMRI* mRNA, polyG peptides, lamin A and C, ubiquitin, SUMO, and p62 protein, while it is rare for polyglutamine (polyQ) and negative for FMRP or TDP-43 [31]. The NIIs are composed of silk-like substances forming a round and membrane-free structure when viewed under electron microscopy. The NIIs are also found in multiple tissues including the peripheral nerves and skin [112].

Neuronal intranuclear inclusion disease (NIID)

Neuronal intranuclear inclusion disease (NIID) is a rare progressive neurodegenerative disease caused by non-coding CGG repeat expansions in the *NOTCH2NLC* gene [18, 43, 100, 110]. As suggestive as the name, it is characterized by ubiquitin- or p62-positive extensive eosinophilic NIIs in central and peripheral nervous tissues, which show similar ultrastructural changes as FXTAS under electron microscopy [60, 101]. However, skin biopsy could show NIIs located in the fibroblast, adipocyte, and epithelial cells of sweat gland ducts, which greatly facilitated the antemortem diagnosis of NIID [102]. The clinical features of NIID show great heterogeneity with combinations of cognitive impairments, stroke-like symptoms, encephalitic episodes, autonomic dysfunction, limb weakness, cerebellar ataxia, parkinsonism, peripheral neuropathy, psychiatric disturbance, visual abnormalities, and other multi-system symptoms [57, 119]. The MRI features in adult-onset NIID include diffuse white matter lesions, DWI and FLAIR hyperintensities in the corpus callosum, and DWI linear hyperintensity along the cortico-medullary junction, which are strikingly similar to those of FXTAS [43, 57].

According to the age of onset, NIID can be classified into infantile-onset, juvenile-onset, and adult-onset subgroups [64]. Based on the relationship between phenotype and genotype, researchers divided familial genetically positive adult-onset NIID cases into three subgroups: parkinsonism-dominant NIID, muscle weakness-dominant NIID, and dementia-dominant NIID [110]. Nevertheless, several studies have indicated that carriers with more than 300 repeats of expanded CGG show a mild or asymptomatic phenotype [20, 127].

Beyond NIID, expanded CGG repeats in *NOTCH2NLC* are occasionally related to a small proportion of Parkinson's disease (PD) [65, 97, 110], multiple system atrophy (MSA) [29], essential tremor (ET) [105], degenerative dementia [4, 101], ALS [46, 130], inherited peripheral neuropathy [118], distal motor neuropathy [122, 128], mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes (MELAS) [57, 123], and oculopharyngodistal myopathy (OPDM) [127]. Intriguingly, CGG expansion in *NOTCH2NLC* was rarely detected in NIID cases reported in people of Caucasian descent, suggesting that NIID is likely to be genetically heterogeneous among different ethnic groups [15].

Oculopharyngeal myopathy with leukoencephalopathy (OPML)

OPML is an extremely rare genetic disease characterized by ptosis, ophthalmoplegia, dysphagia, dysarthria, and limb muscle weakness [43]. It has only been reported in a four-generation Japanese family, in which seven individuals showed a similar but variable clinical phenotype including severe gastrointestinal dysmotility, respiratory failure, ataxia, bladder disturbances, tremor, and dilated cardiomyopathy. Individuals with diffuse limb weakness underwent muscle biopsies, which showed a nonspecific myopathic change. Cerebral MRI in three affected individuals showed brain atrophy with T2 hyperintensity signals in the white matter consistent with leukoencephalopathy. Intriguingly, one patient showed leukoencephalopathy that was strikingly similar to FXTAS and NIID with characteristic DWI changes at the frontal corticomedullary junctions. Inspired by the common clinical characteristics and neuroimaging features, expanded CGG repeats in whole-genome sequence data were directly explored in one individual from the family. Researchers identified and verified heterozygous trinucleotide repeat expansion CGG in the *NUTM2B-AS1* gene on chromosome 10q22 as the causative gene of OPML.

Oculopharyngodistal myopathy (OPDM)

OPDM is a rare, adult-onset inherited neuromuscular disorder characterized by progressive ptosis, external ophthalmoplegia, and weakness of the masseter, facial, pharyngeal, and distal limb muscles [26, 93]. Recent studies indicated that CGG repeat expansions in the 5'UTR of the *LRP12* [43], *GIPC1* [19], *NOTCH2NLC* [127], and *RILPL1* [129] genes were associated with OPDM types 1, 2, 3, and 4, respectively. Clinical manifestations of all OPDM subtypes showed a characteristic distribution of muscle involvement. Moreover, a small part of patients with OPDM1/2/4 had extra-muscular symptoms, while most patients with OPDM3 simultaneously had different

degrees of leukoencephalopathy, peripheral neuropathy, and other neurological manifestations [82, 127, 128]. The myopathological features were myopathic changes of differing severity characterized by the presence of rimmed vacuoles (RVs) and myeloid bodies in the cytoplasm of muscle cells [19, 82]. NIIs, similar to those of FXTAS and NIID, could be observed in muscle specimens or skin biopsy samples in OPDM1-4, which could facilitate the diagnostic workflow [81].

The common spectrum of polyG diseases

Clinical manifestations of polyG diseases

The polyG diseases are adult-onset, slowly progressive, multi-system neurodegenerative disorders that primarily involve the central and peripheral nervous systems and the muscular system. Movement disorder, cognitive disturbance, muscle weakness, and peripheral neuropathy are the common clinical characteristics of polyG diseases.

Movement disorder phenotype

Different types of tremor have been described in patients with FXTAS, including ET-like, rest, and cerebellar tremors, and the different tremor types might coexist. In several studies, patients diagnosed with ET were screened for CGG mutations in *FMRI*, *NOTCH2NLC* [105, 125], and *GIPC1* [28], with positive findings in a few patients. Additionally, some OPML patients exhibited mild ataxia or tremor [43]. Collectively, tremors were the most common symptoms of movement disorders in polyG diseases.

Approximately 29% to 60% of FXTAS patients were misdiagnosed as parkinsonism [48, 80, 92]. Parkinsonism is a series of neurodegenerative syndrome characterized by bradykinesia, resting tremor, rigidity, and loss of postural reflexes. Intermediate-length CGG repeat expansion of *NOTCH2NLC* have been associated with PD that is responsive to small doses of levodopa over many years [65, 97]. Expansion of CGG in *GIPC1* was also found in ten out of 1,036 patients diagnosed with PD [28]. In a study of clinical characteristics of *LRP12*-OPDM, one out of 65 patients presented with idiopathic PD in her 70 s [55]. Therefore, genetic testing of non-coding CGG repeat expansion in patients with movement disorders of unknown etiology may be helpful for accurate diagnosis and effective treatment.

Cognitive deficit phenotypes

Progressive memory loss is a predominant symptom in some patients with FXTAS or NIID [12]. Due to diffuse white matter lesions and DWI hyperintensity along the cortico-medullary junction of frontal-parietal lobes, executive function deficits overlapping with cognitive

decline are common symptoms in patients with FXTAS [12] or NIID. Although no cognitive disturbances were reported in OPML patients, extensive white matter lesions and brain atrophy were observed in cerebral MRIs, still the cognitive function of patients was uncertain due to lack of long-term follow-up [43].

The phenotype of OPDMs mainly presents with myopathy and less involvement of the central nervous system (CNS), while a few patients show CNS manifestations [82]. The largest retrospective study of the disorder included a cohort of 64 individuals who were identified as *LRP12* CGG expansion positive, in which 92% showed no significant CNS manifestations [55]. In this study, only one of five patients who underwent biopsy had intranuclear tubulofilamentous inclusions in the bicep muscle with course over eight years. However, a general autopsy was performed in one OPDM1 patient, who had no reported neurological disorder during his lifetime. Unexpectedly, it revealed the presence of NIIs in almost all organs including CNS and peripheral nerves [90]. Approximately 44% of OPDM2 patients with *GIPC1* CGG repeat showed a cognitive deficit [28], and almost all OPDM3 patients had different white matter changes. Collectively, we hypothesized that patients with polyG-related myopathies initially and mainly showed myopathy-based manifestations, while different degrees of cognitive impairment might be unidentified or occur gradually with disease development.

Autonomic dysfunction

Autonomic dysfunction is commonly noticed in FXTAS patients, especially impotence and frequent dysfunction of bowel and bladder [33, 45]. Autonomic dysfunction is also prominent and common in patients with adult-onset NIID, and unexplained urinary disturbance is sometimes the only symptom preceding other neurological symptoms for many years [77]. Bilateral miosis is a unique manifestation and diagnostic indicator for NIID [64]. Other autonomic disturbances such as gastrointestinal dysfunction, orthostatic hypotension, arrhythmia, and sexual dysfunction are often recorded in the medical history of NIID patients. Gastrointestinal dysmotility and bladder function have been reported in OPML patients, suggesting the possibility that autonomic nerves are affected by NIIs [43]. Collectively, autonomic dysfunctions in polyG diseases are very common and highly heterogeneous, and can occur more than a decade before typical neuroimaging changes, especially in NIID and FXTAS.

Muscle-weakness phenotypes

The names OPDM and OPML indicate that weaknesses of ocular, bulbar, and limb muscles are likely the primary

manifestations of these diseases. Additionally, cardiac muscle can be affected, as with dilated cardiomyopathy observed in NIID, OPML and OPDM1 patients [43, 83, 90]. More than half of familial NIID patients showed a muscle weakness-dominant phenotype, which was usually attributed to neurogenic causes, while myopathic origins should not be ignored in some NIID cases. These muscle weakness-dominant patients usually had a younger age at onset, and had a family history associated with larger-size repeat expansions [101]. Similarly, some FXTAS patients may also present with lower limb muscle weakness [45]. The phenotype of muscle weakness in polyG diseases may be associated with NIIs being widespread in nerves and muscles.

Neuroradiological changes of polyG diseases

Neuroimaging of NIID is a very sensitive and specific biomarker showing that DWI high-intensity signals are distributed along the corticomedullary junction in the

frontal and parietal lobes (Fig. 1A). These signals gradually extend along the corticomedullary junction of the whole cerebrum as the disease progresses, but usually do not expand into the deep white matter even late in disease progression [101]. High-intensity subcortical DWI signals in NIID strongly correlate with pathological spongiotic changes of NIID, but the mechanism is still unclear. The “cockscomb pattern” or “ribbon sign” seen along the corticomedullary junction on DWI are characteristic findings in patients with polyG disorders. However, the high-intensity signal at the corticomedullary junction is often not obvious in the early stage of the disease, and can disappear in some patients [49], and thus diagnosis of such diseases should not be overly dependent on the ribbon sign.

White matter hyperintensities and brain atrophy on MRI are other common imaging changes in polyG diseases, including FXTAS, NIID, OPML, and OPMD2-3. The cerebellar involvement in FXTAS is more obvious

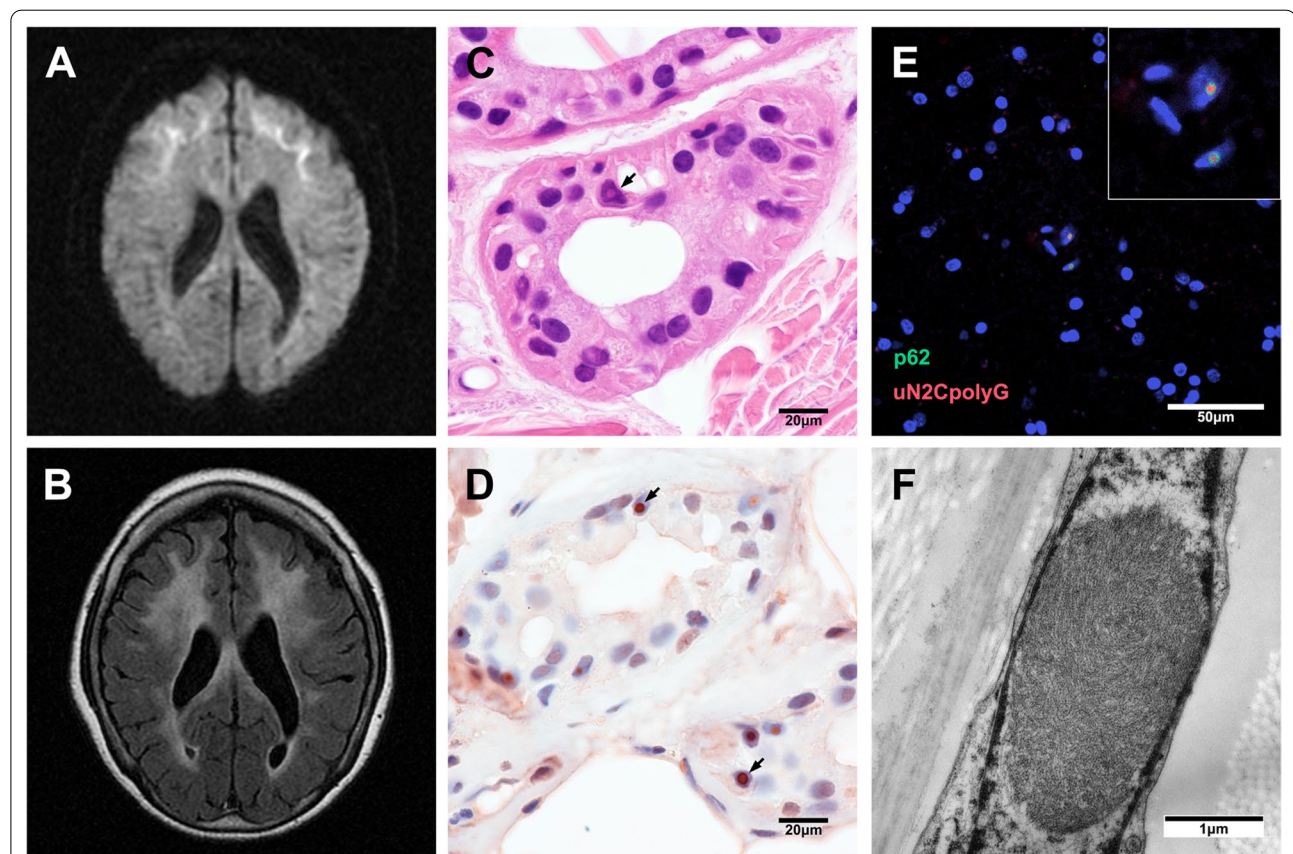


Fig. 1 Typical brain radiological and pathological features in patient with *NOTCH2NLC*-CGG expansion. **A and B** Representative brain DWI image and T2-weighted image of patients with *NOTCH2NLC*-related polyG disease. **C** Representative H&E staining image show eosinophilic NIIs (black arrow) in sweat gland cells of skin **D** Representative immunostaining image show p62-positive NIIs (black arrow) in skin. **E** Representative electron microscopy image revealed Intranuclear tubulofilamentous inclusions in the skin of NIID patient. **F** Immunofluorescence against uN2CpolyG using 4D12 antibody (red, generously providing by Dr. Nicolas Charlet-Berguerand) and p62 antibody (green) on brain section from patient with *NOTCH2NLC*-CGG repeat expansion. Nuclei were counterstained with DAPI

than cerebrum involvement, and MCP sign and general cortical and subcortical atrophy are seen in the majority of males, and in some women [9, 33]. White matter lesions in NIID and OPML are present in the cerebrum, cerebellum, and spinal cord, and most notably in the deep white matter of the brain (Fig. 1B). Leukoencephalopathy is obvious on MRI in 97.4% of sporadic and 76.9% of familial NIID cases [101]. Additionally, persistent DWI hyperintensities in corpus callosum and enhancement of cortical surface were observed in some NIID cases [57, 119]. More than half of FXTAS patients have cerebral microbleeds, and NIIs in the endothelial cells of capillaries, suggesting cerebrovascular dysfunction in FXTAS [91]. Almost all OPDM patients have different white matter hyperintensities but without typical *NOTCH2NLC*-related high-intensity signals at the corticomedullary junction.

Pathological features of polyG diseases

Each polyG disease has its own pathological features, while the presence of NIIs has been demonstrated in all types of polyG disorders, except for OPML. Eosinophilic NIIs are present throughout the central and peripheral nervous systems, and multiple organs in FXTAS, NIID, and OPDM1 [90]. In contrast, owing to the lack of autopsy histopathological analysis, p62-positive NIIs were only found in biopsied skin and skeletal muscle samples in other types of OPDM patients [129]. The NIIs are round and about 1.5–10 μm in diameter, eosinophilic on hematoxylin–eosin (H&E) staining (Fig. 1C), and positive for p62 and polyG (Fig. 1D and E), and ubiquitin, but negative for polyQ on immunostaining. Under electron microscopy, the inclusions are composed of a pile of fibrous substances often with halos around, but without membranes (Fig. 1F).

Although inclusions are also seen in multiple neurodegenerative diseases such as polyQ diseases [24, 86], frontotemporal dementia (FTLD) [44], OPMD [113, 121] and MSA [84], inclusions in polyG diseases have characteristic immune markers and distribution patterns, suggesting the existence of unique pathological mechanisms. NIIs can be detected more than ten years before the onset of NIID symptoms, and are often found in morphologically intact neurons without obvious neuronal loss [72, 76], suggesting that NIIs may be associated with the progression of the disease. p62-positive rimmed vacuoles in muscle fibers are unique pathological changes in OPDM patients. Muscle biopsy specimens from four OPML patients showed nonspecific myopathic changes [43], but more biopsy specimens are needed to confirm whether OPML has NIIs. Additionally, there is a lack of autopsy results from genetically-positive OPML and OPDM2-3 patients, especially from the CNS. Thus, no direct

evidence has demonstrated whether NIIs are present in neurons of OPDM and OPML. Further studies like post-mortem histopathologic analyses are necessary to clarify any extra-muscular organ involvement and associated symptoms that might have been masked by more obvious features relevant to muscle pathology. NIIs involving multiple systems may be responsible for the complex symptoms of polyG diseases.

Epidemiology and genetics

No prospective longitudinal studies of individuals with polyG diseases have been conducted. Some retrospective studies have been conducted on NIID and FXTAS, which showed significant differences in the population distribution of the two disorders [15, 126]. The reported FXTAS cases were mainly in individuals of Caucasian descent, with a few in Asian populations. Conversely, *NOTCH2NLC*-NIID cases were mainly reported in East Asian populations, with no reports from Caucasian populations. Currently, genetically confirmed cases of OPML and OPDM are mostly from China and Japan, and are rare in Caucasian populations. Whether this difference is related to a founder effect is unclear.

In terms of inheritance, polyG diseases are all dominantly inherited, but many patients with polyG diseases show sporadic pattern with no clear family history. Clinical and genetic anticipation are typical features in most STR-related diseases, while anticipation have not observed in patients with polyG diseases currently. Nucleotide repeat instability and unstable transmission of CGG repeats between parents and offspring have been reported in individual FXTAS, NIID, and OPDM families [5, 20]. These characteristics suggest that polyG diseases may have different molecular genetic mechanisms compared to polyQ diseases.

Mechanisms

The puzzle of trinucleotide STR

STRs are a series of small repetitive DNA units consisting of 2–6 nucleotides, covering around 6.77% of the human genome [16]. STRs contribute to genetic diversity, but may be harmful to humans due to repeat-mediated genome instability [37]. Looking back to the past 30 years of studies in repeat expansion disorders [21], polyQ diseases have been characterized by NIIs containing polyQ translated by CAG repeats [59]. Almost 10 years ago, when investigations into the pathogenesis of polyQ diseases was at its beginning stage, one of the seminal questions was how the same mutation present in ORFs of different genes, and thus different proteins with polyQ expansion, could result in different diseases with some common symptoms [79]. Among them, two proteins (androgen receptor and Ataxin-1) with nuclear

localization signal (NLS) in two typical polyQ diseases (KD and SCA1) demonstrate the concept that it is the change in normal function and conformation of the polyQ protein, induced by an expanded polyQ tract, which initiates the pathogenic process [54, 85].

In those days, research on the pathogenesis of FXTAS is at a standstill since the repeat expansion is confined to the 5'UTR of *FMRI* [51], despite the fact that CGG repeat expansion is believed to be essential, as said by Ammar Al-Chalabi when the pathogenic locus of *C9ORF72* was uncertain, "It's like we knew the street, but we didn't know the exact house" [25]. Fortunately, research on DM1, another disease caused by non-coding trinucleotide repeat expansions, has progressed and demonstrated that repeat sequence-mediated RNA toxicity seems to be involved [117]. Unexpectedly, unlike the typical RNA gain-of-function disease, the majority of the mutant RNA is not located in NIIs, and there is little or no reduction of FMRP in FXTAS [9, 109]. Another hypothesis first proposed by Ranum and colleagues suggested that repeat-associated non-AUG (RAN) translation was involved; this was later confirmed by Todd et al. in FXTAS [111, 136]. Since no or very little FMRpolyA is observed in individuals or disease models of FXTAS, the role of FMRpolyG is an attractive alternative. Furthermore, FMRpolyG is indispensable to the formation of NIIs, as well as mediating the CGG repeat associated toxicity [94]. Recently, FXTAS has its companies that consist of a group of emerging diseases (NIID, OPML, and OPDMs) associated with CGG repeats [135]. PolyG positive NIIs are found in NIID and OPDMs, and expression of polyG by embedding CGG repeat sequence in the 5'UTR of *NOTCH2NLC* is pathogenic in cells and animals. Inspired by polyQ diseases, it led us to infer that polyG is the core pathogenesis of these diseases.

Repeat-mediated genome instability

Although polyG diseases are reported to be autosomal dominant [21], the hereditary mode is uncertain due to random expansion or contraction of the repeats [52]. For example, offspring of FXTAS patients can be unaffected or suffer from FXS [5]. Unexpectedly, males carrying large CGG repeat expansions (up to 300) in *NOTCH2NLC* seem to be asymptomatic, though *NOTCH2NLC* mRNA levels decrease as a consequence of hypermethylation around the CGG repeats, displaying a strikingly different prognosis in contrast to FXS patients with full mutation alleles of *FMRI* [133]. However, these individuals may have children with NIID if the repeat number is contracted during spermatogenesis [20]. The inconsistent repeat number is related to genome instability, and possibly an explanation of variable disease

phenotypes considering the length of polyG tracts translated by the repeats [66].

There are several possible factors underlying repeat-mediated genome instability, including dynamic DNA structures, DNA replication and transcription, DNA repair, and chromatin environment [52]. During DNA replication or transcription, single-stranded DNA can be formed, and the repeat tracts tend to incorporate into different secondary structures determined by the sequence itself. In brief, H-DNA is likely to be formed by homopurine/homopyrimidine mirror repeats [87], G4-DNA by GC-rich sequence, imperfect hairpin by inverted DNA repeat [30], and DNA-unwinding elements by AT-rich sequence [61]. These secondary structures impede replication fork progression and transcription, thus leading to formation of a DNA nick or a double-strand break (DSB), and consequent initiation of DNA repair. The repeat number would undergo unpredictable changes during the process. Moreover, the secondary structures could bring epigenetic changes to the surrounding chromatin, further altering gene expression patterns and repeat instability [104].

For CGG repeats, large repeat length is one of the risks of increasing repeat instability [38]. When located in the lagging strand template during replication, CGG repeats are prone to form hairpins or G4 structures and experience repeat contraction [39, 120]. Large CGG repeats may stall the replication fork, leading to DSBs and chromosome fragility [116, 131]. During DNA transcription, a DNA-RNA hybrid named R-loop is formed on a template strand to regulate gene expression and terminate transcription, leaving the expanding complementary part of the DNA duplex single-stranded to form DNA hairpin (S-loop) [63], G4-DNA or hybrid G4-DNA-RNA structure (G-loop) [2], or a triplex structure of an RNA transcript and a single-stranded DNA portion (H-loop) [78]. These unwanted structures increase the instability of CGG repeats. Additionally, various pathways of DNA repair like homologous recombination (HR), end-joining (EJ) pathways, mismatch repair (MMR), base excision repair (BER), and nucleotide excision repair (NER) further increase the risk of genome instability with CGG repeats [52].

RNA toxicity

RNA-binding proteins (RBPs) are important regulating components of gene expression to support cell viability [7]. The RNA gain-of-function hypothesis was first proposed in DM models based on the observation of nuclear foci containing the myotonic dystrophy protein kinase (*DMPK*) transcript and RBPs interacting with expanded repeats [68, 73, 108], and has become increasingly

significant in non-coding repeat expansion disorders. In this hypothesis, RNA foci, generally formed by interactions between secondary structures of repeat RNAs and RBPs, are the hallmark and are found in many repeat expansion diseases, such as amyotrophic lateral sclerosis and frontotemporal dementia caused by *C9ORF72* (C9ALS/FTD), FXTAS, and many SCAs [132]. Though RNA foci are mostly intranuclear, cytoplasmic RNA foci as well as RNA foci at the edge of the nucleus have also been observed [17, 75]. The canonical role of RNA toxicity in pathogenesis is mis-splicing, as confirmed in DM1, CUG repeat RNA sequesters the muscle blind-like (MBNL) proteins and leads to transcriptome-wide spliceopathy [106]. In addition, miRNA misprocessing [87], transcriptional deregulation [11], global translational inhibition mediated by stress granule [41], and alternative polyadenylation sites [6] have also been proposed in recent years.

For FXTAS, the CGG repeat RNA might compromise the function of various RBPs like Pur-alpha, hnRNP2/B1, CUGBP1, Sam68, and Drosha-DGCR8. Overexpression of most of these proteins can rescue the phenotype in CGG *Drosophila* [47, 95, 96, 99]. However, CGG repeats seem to have a more complex role than could be explained by RNA toxicity alone [94, 111]. Compared with diseases mainly driven by RNA gain-of-function, the pathological repeat range in FXTAS is quite different for the size that is particularly short but close to the pathogenic number of other polyG diseases [21]. Using an RNA FISH probe, CGG repeat RNA formed RNA foci in patients diagnosed with NIID and OPDM types 4. Some RBPs were also found to be colocalized with the RNA foci in polyG diseases like Sam68 and MBNL1 [20, 127, 129]. However, considering the limited role of RNA toxicity in FXTAS, further investigation is needed in other polyG diseases.

polyG protein toxicity

In most conditions, protein translation is a canonical process beginning with an AUG start codon and ending with one of three stop codons (UAA, UGA, and UAG). However, the AUG-dependent translation initially seemed unable to explain the translation process of CGG expansion embedded within 5'UTR, so RAN translation, an AUG-independent translation, was first proposed by Ranum and colleagues [136]. Based on this hypothesis, RAN translation of CGG repeats in FXTAS can lead to the theoretical expression of polyAlanine, polyG and polyArginine containing proteins. However, later studies find that CGG repeats translation still can occur by a canonical ribosome scanning mechanism and initiates upstream of the repeats at near cognate start codons, either a GTG or ACG [50, 94], mainly resulting in expression of one

protein, FMRpolyG, as the CGG repeats are in the glycine frame in regard to these GTG or ACG near cognate start codon. Furthermore, antibodies were developed against these putative polyAlanine and polyG proteins and only FMRpolyG was consistently observed in FXTAS brain sections, while the polyAlanine protein was rarely observed or absent [53, 94]. Thus, RAN translation of CGG repeats into polyAlanine or polyArginine is now considered as a minor pathogenic mechanism, if even present, in FXTAS. Similarly, recent reports indicate that translation of the *NOTCH2NLC* CGG repeats into a novel polyG-containing protein occurs through canonical translation initiation at an AUG start codon located upstream of the repeats, with again the CGG repeats in the glycine frame in regards to this AUG start codon [10, 134]. Collectively, both CGG repeat expansion translation in FXTAS and NIID occur through a canonical mechanism of translation, different from the RAN translation proposed by the Ranum group. Intriguingly, an anti-glycine non-specific antibody that may recognize the polyG domain of protein have been found to be colocalized with p62 in NIIs of OPDM3 and 4, indicating a possibility that CGG expansion embedded in the 5'UTR of corresponding genes (*NOTCH2NLC* or *RILPL1*) may be translated into polyG with the same ribosome-dependent translation mechanism [127, 129], but the potential mechanism of OPDM/OPML still needs to be further studied.

Sellier et al. carried out a series of experiments with a fascinating design and provided crucial evidence for discriminating the role of RNA toxicity from protein toxicity [94]. They found that cell models with expanded CGG repeats embedded in the 5'UTR of *FMRI* formed rare RNA foci compared to the expanded CGG repeats without the surrounding *FMRI* sequence. Transgenic mouse models with mutant 5'UTR that express only the CGG RNA remained indistinguishable from control mice. Beyond that, the authors indicated that FMRpolyG first accumulated in the cytoplasm and formed aggregates, and then these aggregates migrated and formed an inclusion within the cell nucleus. Not surprisingly, the phenomenon reappeared in cell models transfected with *NOTCH2NLC* 5'UTR GGC49-GFP (polyG frame), and though there were no fluorescence signals in the cytoplasm, the nuclear polyG inclusions became larger with time. Furthermore, FMRpolyG and uN2CpolyG disrupted both the morphology and function of nuclear lamina, and overexpression of LAP2 rescued neuron death caused by FMRpolyG [134]. According to the above studies in FXTAS and NIID, the significance of polyG in pathogenesis may overwhelm that of toxic RNA, while how much of a role and how polyG plays role in the pathogenesis of polyG diseases remains unclear.

Future directions for CGG repeat expansion

Learning from polyQ diseases, we can foresee the productive future of pathogenic mechanism studies in polyG diseases. Here, according to the investigations of polyQ diseases and FXTAS, we would like to give some preliminary opinions about future research in polyG diseases (Fig. 2).

First, what is the role of host genes in pathogenesis? Although CGG repeat expansions may lead us to understand the diseases with a common pathological change, the clinical spectrum is quite varied among polyG

diseases [135]. With the same CGG repeat expansion located in non-coding regions and similar repeat-length thresholds for disease penetrance, we wonder if different host genes can partly explain the variety of symptoms. Since the CGG repeat is embedded in uORFs, polyG seems to function independently of the downstream protein expressed by the host genes, while the tissue specific expression pattern of the host gene may limit the toxicity of polyG considering the translation efficiency. Due to separation from the host protein, polyG toxicity is totally different from the mechanism in polyQ that is closely

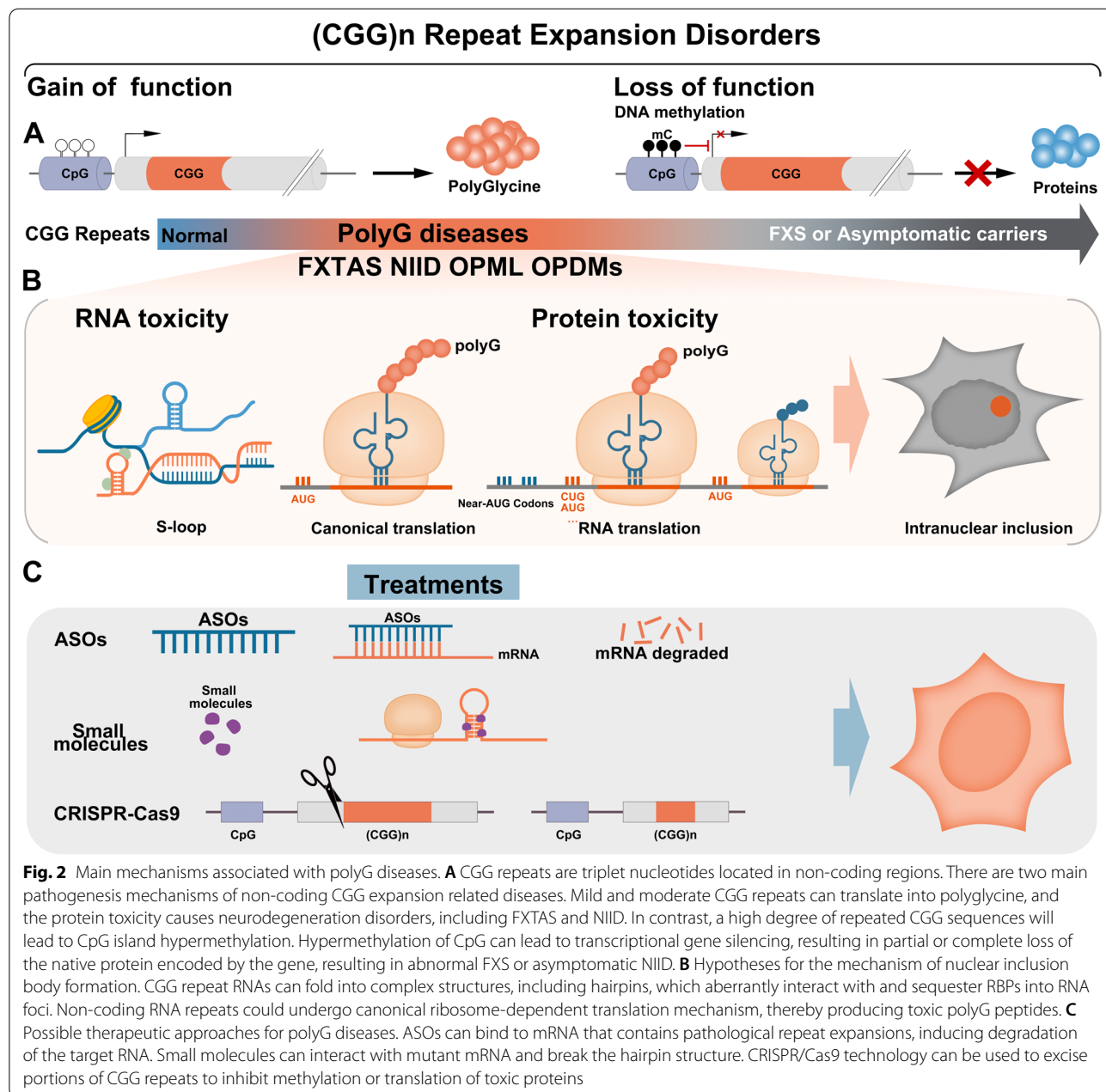


Fig. 2 Main mechanisms associated with polyG diseases. **A** CGG repeats are triplet nucleotides located in non-coding regions. There are two main pathogenesis mechanisms of non-coding CGG expansion related diseases. Mild and moderate CGG repeats can translate into polyglycine, and the protein toxicity causes neurodegeneration disorders, including FXTAS and NIID. In contrast, a high degree of repeated CGG sequences will lead to CpG island hypermethylation. Hypermethylation of CpG can lead to transcriptional gene silencing, resulting in partial or complete loss of the native protein encoded by the gene, resulting in abnormal FXS or asymptomatic NIID. **B** Hypotheses for the mechanism of nuclear inclusion body formation. CGG repeat RNAs can fold into complex structures, including hairpins, which aberrantly interact with and sequester RBPs into RNA foci. Non-coding RNA repeats could undergo canonical ribosome-dependent translation mechanism, thereby producing toxic polyG peptides. **C** Possible therapeutic approaches for polyG diseases. ASOs can bind to mRNA that contains pathological repeat expansions, inducing degradation of the target RNA. Small molecules can interact with mutant mRNA and break the hairpin structure. CRISPR/Cas9 technology can be used to excise portions of CGG repeats to inhibit methylation or translation of toxic proteins

related to the function of the host gene [59]. For example, the way that polyG enters the nucleus is unclear, while polyQ containing protein can be transported to nucleus with NLS expressed by host genes. As shown in previous work, host proteins were partly colocalized with NIIs, especially for *NOTCH2NLC*, the gene which can sometimes be translated into a polyG containing protein [127, 129]. Since CGG repeat translation is initiated upstream of the start codon, it may sequester the translation tools and disrupt the expression of downstream proteins [94]. However, the levels of host proteins seem to be unaffected, and individuals remain asymptomatic even once the expression is down-regulated [20].

Second, what is the role of NIIs? NIIs are formed in a series of triplet expansion diseases but their contribution to the disease itself is uncertain [66, 67]. In polyQ diseases, inclusions mainly form in the nucleus, and may be protective in the early phase because they may isolate the toxic polyQ from binding with functional proteins. Inclusions can be toxic when they turn into irreversible solid aggregates [59]. However, long before the NIIs formed, repeat RNA needs to be translated in the cytoplasm and gradually migrate and settle into the nucleus [94]. As indicated in polyQ diseases, the accepted idea is that the nucleus is an important site of polyQ toxicity partly due to the NLS carried by the host proteins [59]. While in polyG-related diseases, when the host proteins no longer work with the toxic tract [10, 94, 134], would the role of nucleus be more like a final location than an active pathogenic site? Like RVs present in OPDMs [90], would polyG reside in other parts of the cell and have a different pathogenic role? Future advancements in disease research may answer these fascinating questions.

Third, what is the role of polyG? As for CGG repeats located in the 5'UTR of *FMRI* and *NOTCH2NLC*, polyG has been identified to be toxic in different disease models, and it may interact with lamin to disturb nuclear function. Overexpression of LAP2B partly rescues the phenotype of FXTAS animal models [10, 94, 134]. This one affected site cannot provide us a whole picture to fully explain the various clinical outcomes, thus more potential targets need to be found in the future. Although recent knowledge about polyG is limited, the future is not so uncertain since polyQ has been found to be correlated with gene expression, axonal transportation, mitochondrial function, neuronal excitotoxicity, and ion homeostasis [43].

Therapeutic approaches for polyG diseases

There is still a lack of effective treatment for polyG diseases. As more is learned about the pathogenesis of the diseases, several potential therapeutic approaches

targeting the DNA, RNA, and protein levels have been developed [124].

Antisense oligonucleotides (ASOs) are chemically synthesized single stranded nucleic acids that can bind RNA targets and prevent them from forming secondary structures and sequestering the RBPs. Furthermore, the target RNAs can be degraded by RNaseH [8]. ASOs have been successfully applied in treating neurodegenerative diseases caused by repeat expansions. Recent achievements witnessed that ALS fibroblast or induced pluripotent stem cell (iPSC)-derived neurons were rescued by reducing formation of RNA foci and glutamate excitotoxicity [1, 8]. Clinical trials have been conducted in patients with ALS and Huntington's disease (HD) and partly relieve their symptoms [13, 74, 107]. Encouragingly, this kind of therapy has been investigated in models of FXTAS, and specific ASOs greatly improved the clinical and pathological phenotypes in FXTAS mice [22].

Another intervention targeting RNA toxicity is RNA interference (RNAi) strategies, to which purpose is degrading target mRNA and reducing protein expression with RNA molecules including microRNA (miRNA), small interfering RNA (siRNA), and short hairpin RNA (shRNA) are now the most common RNA molecules for RNAi [14]. For C9ALS/FTD, siRNA effectively reduced GGGGCC repeat-containing transcripts and RNA foci formation in cells and a mouse model [40, 69, 70]. For HD, miRNA has been applied and successfully reduced transcript of huntingtin gene (HTT) levels, leading to improvement in neuropathology [75]. For FXTAS, evidence has shown that using siRNA against RNA of a specific retrotransposon, gypsy, is able to modulate neurodegeneration in *Drosophila* [42].

As for intervention in RNA levels, small molecules can also interact with CGG repeat RNAs. 9-hydroxy-5,11-dimethyl-2-(2-(piperidin-1-yl)ethyl)-6H-pyrido[4, 3-b]carbazol-2-ium, can bind CGG repeats in vitro, improve FXTAS-associated splicing defects, and reduce the size and number of pathologic protein aggregates. Other small molecules identified to interact with CGG repeats include phospholipase A2 inhibitors [88], naphthyridine carbamate dimer [36], piperine, geldanamycin [115], and spironolactone [58].

Repairing DNA levels may be important, but such intervention is risky because it may cause higher repeat associated instability, and thus lead to genome mutations or further expansion of the repeat itself. A study suggests that the expanded repeats can be excised from DNA by CRISPR/Cas9 technology [3]. Expression of *FMRI* can be reactivated by using CRISPR/Cas9 to edit *FMRI* full mutation allele (CGG repeats > 200) in FXS iPSCs [62].

Conclusions

Four years ago, following the simple but practical principle, with great comparability in phenotypes leading to great comparability in genes, Ishiura et al. [43], made a breakthrough in STR-related neurodegenerative diseases. They determined that expanded CGG repeats located in non-coding regions were the cause of NIID, OPML, and OPDM type 1. Subsequently, *NOTCH2NLC*-related disorders and other types of OPDM due to abnormal expansion of CGG repeats were reported. Regarding the essential role of polyG in FXTAS and NIID, and polyG-positive NIIs may be also presented in OPDM type 3/4, the diseases caused by CGG repeat expansion can be classified as polyG diseases, highlighting the likely pathogenic role of toxic polyG.

As discussed in this review, the clinical spectrum of polyG diseases is variable, but mainly involves the central and peripheral nervous systems and the muscular system. The disease spectrum may be further enlarged in the future due to the increasing attention on CGG repeat expansion and the advancement of DNA sequencing methods. However, a well-explained polyG-related pathogenesis underlying these different disease phenotypes would be challenging to propose. Until now, the basic research of polyG diseases has focused mainly on modeling the formation and toxicity of polyG in cells and animals, and only one downstream target (nuclear lamina) was discovered [94]. Thus, targeted interventions for the polyG diseases have been stuck on RNAi, gene editing, and small molecules interacting with repeat RNAs, which may lead to genome instability. Future investigations of the diseases may reveal more effective downstream targets of the polyG proteins, and provide us with a more complete picture to better understand, diagnose, and treat the diseases.

Abbreviations

FXTAS: Fragile X-associated tremor/ataxia syndrome; 5'UTR: 5'Untranslated region; NIID: Neuronal intranuclear inclusion disease; OPML: Oculopharyngeal myopathy with leukoencephalopathy; OPDM: Oculopharyngodistal myopathy; uORF: Upstream open reading frame; polyG: Polyglycine; STRs: Short tandem repeats; DM: Myotonic dystrophies; SCAs: Spinocerebellar ataxias; ALS: Amyotrophic lateral sclerosis; KD: Kennedy disease; LRS: Long-read sequencing; NIIs: Intranuclear inclusions; FMR1: Fragile X mental retardation 1; FMRP: FMR1 protein; FXS: Fragile X syndrome; FLAIR: Fluid attenuated inversion recovery; MCP: Middle cerebellar peduncle; DWI: Diffusion-weighted imaging; polyQ: Polyglutamine; PD: Parkinson's disease; MSA: Multiple system atrophy; ET: Essential tremor; MELAS: Mitochondrialencephalomyopathy,lacticacidosis, and stroke-like episodes; RVs: Rimmed vacuoles; CNS: Central nervous system; H&E: Hematoxylin–eosin; FTLD: Frontotemporal dementia; NLS: Nuclear localization signal; RAN: Repeat-associated non-AUG; DSB: Double-strand break; HR: Homologous replication; EJ: End-joining; MMR: Mismatch repair; BER: Base excision repair; NER: Nucleotide excision repair; RBPs: RNA-binding proteins; MBNL: Muscleblind-like; FMRpolyG: PolyGlycine translated by CGG repeat transcripts embedded in the 5'UTR of *FMR1*; uN2CpolyG: Polyglycine translated by CGG repeat transcripts in the 5'UTR of *NOTCH2NLC*; R-loop: DNA-RNA

hybrid; S-loop: R-loop with DNA hairpin; G-loop: R-loop with G4-DNA or hybrid G4-DNA-RNA structure; H-loop: R-loop with a triplex structure of an RNA transcript and a single-stranded DNA portion; ASOs: Antisense oligonucleotides; iPSC: Induced pluripotent stem cell; RNAi: RNA interference; miRNA: MicroRNA; siRNA: Small interfering RNA; shRNA: Short hairpin RNA; DMPK: Myotonic dystrophy protein kinase; AD: Alzheimer's disease; AGD: Argyrophilic grain disease; TDP-43: TAR DNA binding protein 43.

Author contributions

LFT and ZY drafted the manuscript. ZY draw the illustration. DJ, HD and WZ provide critical feedback and revise the manuscript. DJ and HD design this review (conceptualization of the idea), gives an interpretation of data, and writes the article. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (Grants,81460199, 81571219, 82071409, 82171846, 82160252, and U20A20356), Science and technology project of Jiangxi Health Commission (202110028), Double thousand talents program of Jiangxi province (jxsq2019101021), Peking University Medicine Fund of Fostering, Young Scholars'Scientific & Technological Innovation(BMU2021PY003), and Capital's Funds for Health Improvement and Research (2022–4-40716).

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Declarations

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Author details

¹Department of Neurology, Peking University First Hospital, Beijing, China.

²Department of Neurology, The First Affiliated Hospital of Nanchang University, Nanchang, China. ³Beijing Key Laboratory of Neurovascular Disease Discovery, Beijing, China. ⁴Department of Medical Genetics, The First Affiliated Hospital of Nanchang University, Nanchang, China.

Received: 1 April 2022 Accepted: 16 May 2022

Published online: 31 May 2022

References

- Abati E, Bresolin N, Comi G, Corti S (2020) Silence superoxide dismutase 1 (SOD1): a promising therapeutic target for amyotrophic lateral sclerosis (ALS). *Expert Opin Ther Targets* 24:295–310. <https://doi.org/10.1080/14728222.2020.1738390>
- Abu Diab M, Mor-Shaked H, Cohen E, Cohen-Hadad Y, Ram O, Epsztejn-Litman S, Eiges R (2018) The G-rich repeats in and loci are hotspots for local unpairing of DNA. *Genetics* 210:1239–1252. <https://doi.org/10.1534/genetics.118.301672>
- Agtmaal EV, André L, Willemsse M, Cumming SA, Kessel IV, van den Broek WJAA, Gourdon G, Furling D, Mouly V, Monckton DG (2017) CRISPR/Cas9-Induced (CTGCAG)_n repeat instability in the myotonic dystrophy type 1 locus: implications for therapeutic genome editing. *Mol Ther* 25:24–43. <https://doi.org/10.1016/j.ymthe.2016.10.014>
- Araki K, Sone J, Fujioka Y, Masuda M, Ohdake R, Tanaka Y, Nakamura T, Watanabe H, Sobue G (2016) Memory loss and frontal cognitive dysfunction in a patient with adult-onset neuronal intranuclear inclusion disease. *Intern Med* 55:2281–2284. <https://doi.org/10.2169/intermalmedicine.55.5544>

5. Bartlett E, Archibald AD, Francis D, Ling L, Thomas R, Chandler G, Ward L, O'Farrell G, Pandelache A, Delatycki MB et al (2021) Paternal retraction of a fragile X allele to normal size, showing normal function over two generations. *Am J Med Genet A*. <https://doi.org/10.1002/ajmg.a.62500>
6. Batra R, Charizanis K, Manchanda M, Mohan A, Li M, Finn DJ, Goodwin M, Zhang C, Sobczak K, Thornton CA et al (2014) Loss of MBNL leads to disruption of developmentally regulated alternative polyadenylation in RNA-mediated disease. *Mol Cell* 56:311–322. <https://doi.org/10.1016/j.molcel.2014.08.027>
7. Belzil VV, Gendron TF, Petrucelli L (2013) RNA-mediated toxicity in neurodegenerative disease. *Mol Cell Neurosci* 56:406–419. <https://doi.org/10.1016/j.mcn.2012.12.006>
8. Bennett C (2019) Therapeutic antisense oligonucleotides are coming of age. *Annu Rev Med* 70:307–321. <https://doi.org/10.1146/annurev-med-041217-010829>
9. Berry-Kravis E, Abrams L, Coffey SM, Hall DA, Greco C, Gane LW, Grigsby J, Bourgeois JA, Finucane B, Jacquemont S et al (2007) Fragile X-associated tremor/ataxia syndrome: clinical features, genetics, and testing guidelines. *Mov Disord* 22:2018–2030. <https://doi.org/10.1002/mds.21493> (quiz 2140)
10. Boivin M, Deng J, Pfister V, Grandgirard E, Oulad-Abdelghani M, Morlet B, Ruffenach F, Negroni L, Koebel P, Jacob H et al (2021) Translation of GGC repeat expansions into a toxic polyglycine protein in NIID defines a novel class of human genetic disorders: The polyG diseases. *Neuron* 109(1825–1835):e1825. <https://doi.org/10.1016/j.neuron.2021.03.038>
11. Botta A, Vallo L, Rinaldi F, Bonifazi E, Amati F, Biancolella M, Gambardella S, Mancinelli E, Angelini C, Meola G et al (2007) Gene expression analysis in myotonic dystrophy: indications for a common molecular pathogenic pathway in DM1 and DM2. *Gene Expr* 13:339–351. <https://doi.org/10.3727/000000006781510705>
12. Cabal-Herrera AM, Tassanakijpanich N, Salcedo-Arellano MJ, Hagerman RJ (2020) Fragile X-associated tremor/ataxia syndrome (FXTAS): pathophysiology and clinical implications. *Int J Mol Sci*. <https://doi.org/10.3390/ijms21124391>
13. Cappella M, Pradat PF, Querin G, Biferi MG (2020) Beyond the traditional clinical trials for amyotrophic lateral sclerosis and the future impact of gene therapy. *J Neuromuscul Dis* 8:1–14. <https://doi.org/10.3233/jnd-200531>
14. Castanotto D, Rossi JJ (2009) The promises and pitfalls of RNA-interference-based therapeutics. *Nature* 457:426–433. <https://doi.org/10.1038/nature07758>
15. Chen Z, Yan Yau W, Jaunmuktane Z, Tucci A, Sivakumar P, Gagliano Taliun SA, Turner C, Efthymiou S, Ibanez K, Sullivan R et al (2020) Neuronal intranuclear inclusion disease is genetically heterogeneous. *Ann Clin Transl Neurol* 7:1716–1725. <https://doi.org/10.1002/acn3.51151>
16. Chintalaphani SR, Pineda SS, Deveson IW, Kumar KR (2021) An update on the neurological short tandem repeat expansion disorders and the emergence of long-read sequencing diagnostics. *Acta Neuropathol Commun* 9:98. <https://doi.org/10.1186/s40478-021-01201-x>
17. Cooper-Knock J, Higginbottom A, Stopford MJ, Highley JR, Ince PG, Wharton SB, Pickering-Brown S, Kirby J, Hautbergue GM, Shaw PJ (2015) Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy. *Acta Neuropathol* 130:63–75. <https://doi.org/10.1007/s00401-015-1429-9>
18. Deng J, Gu M, Miao Y, Yao S, Zhu M, Fang P, Yu X, Li P, Su Y, Huang J et al (2019) Long-read sequencing identified repeat expansions in the 5'UTR of the NOTCH2NLC gene from Chinese patients with neuronal intranuclear inclusion disease. *J Med Genet* 56:758–764. <https://doi.org/10.1136/jmedgenet-2019-106268>
19. Deng J, Yu J, Li P, Luan X, Cao L, Zhao J, Yu M, Zhang W, Lv H, Xie Z et al (2020) Expansion of GGC repeat in GIPC1 is associated with oculopharyngodistal myopathy. *Am J Hum Genet* 106:793–804. <https://doi.org/10.1016/j.ajhg.2020.04.011>
20. Deng J, Zhou B, Yu J, Han X, Fu J, Li X, Xie X, Zhu M, Zheng Y, Guo X et al (2021) Genetic origin of sporadic cases and RNA toxicity in neuronal intranuclear inclusion disease. *J Med Genet*. <https://doi.org/10.1136/jmedgenet-2020-107649>
21. Depienne C, Mandel J-L (2021) 30 years of repeat expansion disorders: What have we learned and what are the remaining challenges? *Am J Hum Genet* 108:764–785. <https://doi.org/10.1016/j.ajhg.2021.03.011>
22. Derbis M, Kul E, Niewiadomska D, Sekrecki M, Sobczak K (2021) Short antisense oligonucleotides alleviate the pleiotropic toxicity of RNA harboring expanded CGG repeats. *Nat Commun* 12:1265. <https://doi.org/10.1038/s41467-021-21021-w>
23. Devys D, Lutz Y, Rouyer N, Bellocq JP, Mandel JL (1993) The FMR-1 protein is cytoplasmic, most abundant in neurons and appears normal in carriers of a fragile X premutation. *Nat Genet* 4:335–340. <https://doi.org/10.1038/ng0893-335>
24. DiFiglia M, Sapp E, Chase KO, Davies SW, Bates GP, Vonsattel JP, Aronin N (1997) Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* 277:1990–1993. <https://doi.org/10.1126/science.277.5334.1990>
25. Dolgin E (2017) Genetics: the hexanucleotide hex. *Nature* 550:S106–S108. <https://doi.org/10.1038/550S106a>
26. Durmus H, Laval SH, Deymeer F, Parman Y, Kiyan E, Gokyigit M, Ertekin C, Ercan I, Solakoglu S, Karcagi V et al (2011) Oculopharyngodistal myopathy is a distinct entity: clinical and genetic features of 47 patients. *Neurology* 76:227–235. <https://doi.org/10.1212/WNL.0b013e318207b043>
27. Ehrlich ME, Ellerby LM (2021) Neuronal intranuclear inclusion disease: polyglycine protein is the culprit. *Neuron* 109:1757–1760. <https://doi.org/10.1016/j.neuron.2021.05.018>
28. Fan Y, Shen S, Yang J, Yao D, Li M, Mao C, Wang Y, Hao X, Ma D, Li J et al (2022) GIPC1 CGG repeat expansion is associated with movement disorders. *Ann Neurol*. <https://doi.org/10.1002/ana.26325>
29. Fang P, Yu Y, Yao S, Chen S, Zhu M, Chen Y, Zou K, Wang L, Wang H, Xin L et al (2020) Repeat expansion scanning of the NOTCH2NLC gene in patients with multiple system atrophy. *Ann Clin Transl Neurol* 7:517–526. <https://doi.org/10.1002/acn3.51021>
30. Gacy AM, Goellner G, Juranić N, Macura S, McMurray CT (1995) Trinucleotide repeats that expand in human disease form hairpin structures in vitro. *Cell* 81:533–540. [https://doi.org/10.1016/0092-8674\(95\)90074-8](https://doi.org/10.1016/0092-8674(95)90074-8)
31. Gelpi E, Botta-Orfila T, Bodi L, Marti S, Kovacs G, Grau-Rivera O, Lozano M, Sanchez-Valle R, Munoz E, Valdeoriola F et al (2017) Neuronal intranuclear (hyaline) inclusion disease and fragile X-associated tremor/ataxia syndrome: a morphological and molecular dilemma. *Brain* 140:e51. <https://doi.org/10.1093/brain/awx156>
32. Hagerman PJ, Hagerman RJ (2004) Fragile X-associated tremor/ataxia syndrome (FXTAS). *Ment Retard Dev Disabil Res Rev* 10:25–30. <https://doi.org/10.1002/mrdd.20005>
33. Hagerman RJ, Hagerman P (2016) Fragile X-associated tremor/ataxia syndrome - features, mechanisms and management. *Nat Rev Neurol* 12:403–412. <https://doi.org/10.1038/nrneurol.2016.82>
34. Hagerman RJ, Leavitt BR, Farzin F, Jacquemont S, Greco CM, Brunberg JA, Tassone F, Hessler D, Harris SW, Zhang L et al (2004) Fragile-X-associated tremor/ataxia syndrome (FXTAS) in females with the FMR1 premutation. *Am J Hum Genet* 74:1051–1056. <https://doi.org/10.1086/420700>
35. Hagerman RJ, Leehey M, Heinrichs W, Tassone F, Wilson R, Hills J, Grigsby J, Gage B, Hagerman PJ (2001) Intention tremor, parkinsonism, and generalized brain atrophy in male carriers of fragile X. *Neurology* 57:127–130. <https://doi.org/10.1212/wnl.57.1.127>
36. Hagihara M, He H, Kimura M, Nakatani K (2012) A small molecule regulates hairpin structures in d(CGG) trinucleotide repeats. *Bioorg Med Chem Lett* 22:2000–2003. <https://doi.org/10.1016/j.bmcl.2012.01.030>
37. Hannan AJ (2018) Tandem repeats mediating genetic plasticity in health and disease. *Nat Rev Genet* 19:286–298. <https://doi.org/10.1038/nrg.2017.115>
38. Hansen RS, Canfield TK, Lamb MM, Gartler SM, Laird CD (1993) Association of fragile X syndrome with delayed replication of the FMR1 gene. *Cell* 73:1403–1409. [https://doi.org/10.1016/0092-8674\(93\)90365-w](https://doi.org/10.1016/0092-8674(93)90365-w)
39. Hirst MC, White PJ (1998) Cloned human FMR1 trinucleotide repeats exhibit a length- and orientation-dependent instability suggestive of in vivo lagging strand secondary structure. *Nucleic Acids Res* 26:2353–2358. <https://doi.org/10.1093/nar/26.10.2353>
40. Hu J, Rigo F, Prakash TP, Corey DR (2016) Recognition of c9orf72 mutant RNA by single-stranded silencing RNAs. *Nucleic Acid Ther* 27:87–94. <https://doi.org/10.1089/nat.2016.0655>
41. Huichalaf C, Sakai K, Jin B, Jones K, Wang G-L, Schoser B, Schneider-Gold C, Sarkar P, Pereira-Smith OM, Timchenko N et al (2010) Expansion of

- CUG RNA repeats causes stress and inhibition of translation in myotonic dystrophy 1 (DM1) cells. *FASEB J* 24:3706–3719. <https://doi.org/10.1096/fj.09-151159>
42. Tan H, Qurashi A, Poidevin M, Nelson D (2012) Retrotransposon activation contributes to fragile X premutation rCGG-mediated neurodegeneration. *Hum Mol Genet* 21(1):57–65. <https://doi.org/10.1093/hmg/ddr437>
 43. Ishiura H, Shibata S, Yoshimura J, Suzuki Y, Qu W, Doi K, Almansour MA, Kikuchi JK, Taira M, Mitsui J et al (2019) Noncoding CGG repeat expansions in neuronal intranuclear inclusion disease, oculopharyngodistal myopathy and an overlapping disease. *Nat Genet* 51:1222–1232. <https://doi.org/10.1038/s41588-019-0458-z>
 44. Jackson M, Lowe J (1996) The new neuropathology of degenerative frontotemporal dementias. *Acta Neuropathol* 91:127–134. <https://doi.org/10.1007/s004010050403>
 45. Jacquemont S, Hagerman RJ, Leehey M, Grigsby J, Zhang L, Brunberg JA, Greco C, Des Portes V, Jardini T, Levine R et al (2003) Fragile X premutation tremor/ataxia syndrome: molecular, clinical, and neuroimaging correlates. *Am J Hum Genet* 72:869–878. <https://doi.org/10.1086/374321>
 46. Jih KY, Chou YT, Tsai PC, Liao YC, Lee YC (2021) Analysis of NOTCH2NLCCGG repeat expansion in Taiwanese patients with amyotrophic lateral sclerosis. *Neurobiol Aging*. <https://doi.org/10.1016/j.neurobiolaging.2021.07.011>
 47. Jin P, Duan R, Qurashi A, Qin Y, Tian D, Rosser TC, Liu H, Feng Y, Warren ST (2007) Pur alpha binds to rCGG repeats and modulates repeat-mediated neurodegeneration in a Drosophila model of fragile X tremor/ataxia syndrome. *Neuron* 55:556–564. <https://doi.org/10.1016/j.neuron.2007.07.020>
 48. Juncos JL, Lazarus JT, Graves-Allen E, Shubeck L, Rusin M, Novak G, Hamilton D, Rohr J, Sherman SL (2011) New clinical findings in the fragile X-associated tremor ataxia syndrome (FXTAS). *Neurogenetics* 12:123–135. <https://doi.org/10.1007/s10048-010-0270-5>
 49. Kawarabayashi T, Nakamura T, Seino Y, Hirohata M, Mori F, Wakabayashi K, Ono S, Harigaya Y, Shoji M (2018) Disappearance of MRI imaging signals in a patient with neuronal intranuclear inclusion disease. *J Neurol Sci* 388:1–3. <https://doi.org/10.1016/j.jns.2018.02.038>
 50. Kearse MG, Green KM, Krans A, Rodriguez CM, Linsalata AE, Goldstroh AC, Todd PK (2016) CGG repeat-associated non-AUG translation utilizes a cap-dependent scanning mechanism of initiation to produce toxic proteins. *Mol Cell* 62:314–322. <https://doi.org/10.1016/j.molcel.2016.02.034>
 51. Kenneson A, Zhang F, Hagedorn CH, Warren ST (2001) Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. *Hum Mol Genet* 10:1449–1454. <https://doi.org/10.1093/hmg/10.14.1449>
 52. Khristich AN, Mirkin SM (2020) On the wrong DNA track: molecular mechanisms of repeat-mediated genome instability. *J Biol Chem* 295:4134–4170. <https://doi.org/10.1074/jbc.REV119.007678>
 53. Krans A, Kearse MG, Todd PK (2016) Repeat-associated non-AUG translation from antisense CCG repeats in fragile X tremor/ataxia syndrome. *Ann Neurol* 80:871–881. <https://doi.org/10.1002/ana.24800>
 54. Kratter IH, Finkbeiner S (2010) PolyQ disease: too many Qs, too much function? *Neuron* 67:897–899. <https://doi.org/10.1016/j.neuron.2010.09.012>
 55. Kumutponpanich T, Ogasawara M, Ozaki A, Ishiura H, Tsuji S, Minami N, Hayashi S, Noguchi S, Iida A, Nishino I et al (2021) Clinicopathologic features of oculopharyngodistal myopathy with LRP12 CGG repeat expansions compared with other oculopharyngodistal myopathy subtypes. *JAMA Neurol* 78:853–863. <https://doi.org/10.1001/jamaneurol.2021.1509>
 56. La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH (1991) Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 352:77–79. <https://doi.org/10.1038/352077a0>
 57. Liang H, Wang B, Li Q, Deng J, Wang L, Wang H, Li X, Zhu M, Cai Y, Wang Z et al (2020) Clinical and pathological features in adult-onset NIID patients with cortical enhancement. *J Neurol* 267:3187–3198. <https://doi.org/10.1007/s00415-020-09945-7>
 58. Licata N, Cristofani R, Salomonsson S, Wilson K, Kempthorne L, Vaizoglu D, D'Agostino V, Pollini D, Loffredo R, Pancher M et al (2022) C9orf72 ALS/FTD dipeptide repeat protein levels are reduced by small molecules that inhibit PKA or enhance protein degradation. *EMBO J* 41:e105026. <https://doi.org/10.15252/embj.2020105026>
 59. Lieberman AP, Shakkottai VG, Albin RL (2019) Polyglutamine repeats in neurodegenerative diseases. *Annu Rev Pathol*. <https://doi.org/10.1146/annurev-pathmechdis-012418-012857>
 60. Lindenberg R, Rubinstein LJ, Herman MM, Haydon GB (1968) A light and electron microscopy study of an unusual widespread nuclear inclusion body disease. A possible residuum of an old herpesvirus infection. *Acta Neuropathol* 10:54–73. <https://doi.org/10.1007/bf00690510>
 61. Liu G, Bissler J, Sinden R, Leffak M (2007) Unstable spinocerebellar ataxia type 10 (ATTCT*(AGAAT) repeats are associated with aberrant replication at the ATX10 locus and replication origin-dependent expansion at an ectopic site in human cells. *Mol Cell Biol* 27:7828–7838. <https://doi.org/10.1128/mcb.01276-07>
 62. Liu XS, Wu H, Krzisch M, Wu X, Graef J, Muffat J, Hnisz D, Li CH, Yuan B, Xu C (2018) Rescue of fragile X syndrome neurons by DNA methylation editing of the FMR1 gene. *Cell* 172(5):979–992. <https://doi.org/10.1016/j.cell.2018.01.012>
 63. Loomis EW, Sanz LA, Chédin F, Hagerman PJ (2014) Transcription-associated R-loop formation across the human FMR1 CGG-repeat region. *PLoS Genet* 10:e1004294. <https://doi.org/10.1371/journal.pgen.1004294>
 64. Lu X, Hong D (2021) Neuronal intranuclear inclusion disease: recognition and update. *J Neural Transm (Vienna)* 128:295–303. <https://doi.org/10.1007/s00702-021-02313-3>
 65. Ma D, Tan YJ, Ng ASL, Ong HL, Sim W, Lim WK, Teo JX, Ng EYL, Lim EC, Lim EW et al (2020) Association of NOTCH2NLCCGG repeat expansions with parkinson disease. *JAMA Neurol* 77:1559–1563. <https://doi.org/10.1001/jamaneurol.2020.3023>
 66. Malik I, Kelley CP, Wang ET, Todd PK (2021) Molecular mechanisms underlying nucleotide repeat expansion disorders. *Nat Rev Mol Cell Biol* 22:589–607. <https://doi.org/10.1038/s41580-021-00382-6>
 67. Malik I, Tseng YJ, Wright SE, Zheng K, Ramaiyer P, Green KM, Todd PK (2021) SRSF protein kinase 1 modulates RAN translation and suppresses CGG repeat toxicity. *EMBO Mol Med* 13:e14163. <https://doi.org/10.15252/emmm.202114163>
 68. Mankodi A, Urbinati CR, Yuan QP, Moxley RT, Sansone V, Krym M, Henderson D, Schalling M, Swanson MS, Thornton CA (2001) Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum Mol Genet* 10:2165–2170. <https://doi.org/10.1093/hmg/10.19.2165>
 69. Martier R, Liefhebber JM, García-Osta A, Miniarikova J, Cuadrado-Tejedor M, Espelosin M, Ursua S, Petry H, Deventer SV, Evers MM (2019) Targeting RNA-mediated toxicity in C9ORF72 ALS/FTD by RNAi-based gene therapy. *Mol Ther-Nucleic Acids* 16:26–37. <https://doi.org/10.1016/j.omtn.2019.02.001>
 70. Martier R, Liefhebber JM, Miniarikova J, Tom V, Snapper J, Kolder I, Petry H, Deventer SV, Evers MM, Konstantinova P (2019) Artificial microRNAs targeting C9ORF72 have the potential to reduce accumulation of the intra-nuclear transcripts in ALS and FTD patients. *Mol Ther-Nucleic Acids* 14:593–608. <https://doi.org/10.1016/j.omtn.2019.01.010>
 71. Martin EM, Zhu Y, Kraan CM, Kumar KR, Godler DE, Field M (2021) Men with premutation alleles of less than 71 CGG repeats have low risk of being affected with fragile X-associated tremor/ataxia syndrome (FXTAS). *J Med Genet*. <https://doi.org/10.1136/jmedgenet-2021-107758>
 72. McFadden K, Hamilton RL, Insalaco SJ, Lavine L, Al-Mateen M, Wang G, Wiley CA (2005) Neuronal intranuclear inclusion disease without polyglutamine inclusions in a child. *J Neuropathol Exp Neurol* 64:545–552. <https://doi.org/10.1093/jnen/64.6.545>
 73. Miller JW, Urbinati CR, Teng-Umnuy P, Stenberg MG, Byrne BJ, Thornton CA, Swanson MS (2000) Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *EMBO J* 19:4439–4448
 74. Miller TM, Pestronk A, David W, Rothstein J, Simpson E, Appel SH, Andres PL, Mahoney K, Allred P, Alexander K (2013) An antisense oligonucleotide against SOD1 delivered intrathecally for patients with SOD1 familial amyotrophic lateral sclerosis: a phase 1, randomised, first-in-man study. *The Lancet Neurol* 12:435–442
 75. Mizielinska S, Lashley T, Norona FE, Clayton EL, Ridler CE, Fratta P, Isaacs AM (2013) C9orf72 frontotemporal lobar degeneration is characterised

- by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol* 126:845–857. <https://doi.org/10.1007/s00401-013-1200-z>
76. Motoki M, Nakajima H, Sato T, Tada M, Kakita A, Arawaka S (2018) Neuronal intranuclear inclusion disease showing intranuclear inclusions in renal biopsy 12 years earlier. *Neurology* 91:884–886. <https://doi.org/10.1212/WNL.0000000000006480>
 77. Nakamura M, Ueki S, Kubo M, Yagi H, Sasaki R, Okada Y, Akiguchi I, Kusaka H, Kondo T (2018) Two cases of sporadic adult-onset neuronal intranuclear inclusion disease preceded by urinary disturbance for many years. *J Neurol Sci* 392:89–93. <https://doi.org/10.1016/j.jns.2018.07.012>
 78. Neil AJ, Liang MU, Khristich AN, Shah KA, Mirkin SM (2018) RNA-DNA hybrids promote the expansion of Friedreich's ataxia (GAA)_n repeats via break-induced replication. *Nucleic Acids Res* 46:3487–3497. <https://doi.org/10.1093/nar/gky099>
 79. Nelson DL, Orr HT, Warren ST (2013) The unstable repeats—three evolving faces of neurological disease. *Neuron* 77:825–843. <https://doi.org/10.1016/j.neuron.2013.02.022>
 80. Niu YQ, Yang JC, Hall DA, Leehey MA, Tassone F, Olichney JM, Hagerman RJ, Zhang L (2014) Parkinsonism in fragile X-associated tremor/ataxia syndrome (FXTAS): revisited. *Parkinsonism Relat Disord* 20:456–459. <https://doi.org/10.1016/j.parkreldis.2014.01.006>
 81. Ogasawara M, Eura N, Nagaoka U, Sato T, Arahata H, Hayashi T, Okamoto T, Takahashi Y, Mori-Yoshimura M, Oya Y et al (2021) Intranuclear inclusions in skin biopsies are not limited to neuronal intranuclear inclusion disease but can also be seen in oculopharyngodistal myopathy. *Neuropathol Appl Neurobiol*. <https://doi.org/10.1111/nan.12787>
 82. Ogasawara M, Iida A, Kumutpongpanich T, Ozaki A, Oya Y, Konishi H, Nakamura A, Abe R, Takai H, Hanajima R et al (2020) CGG expansion in NOTCH2NLC is associated with oculopharyngodistal myopathy with neurological manifestations. *Acta Neuropathol Commun* 8:204. <https://doi.org/10.1186/s40478-020-01084-4>
 83. Oyer CE, Cortez S, O'Shea P, Popovic M (1991) Cardiomyopathy and myocyte intranuclear inclusions in neuronal intranuclear inclusion disease: a case report. *Hum Pathol* 22:722–724. [https://doi.org/10.1016/0046-8177\(91\)90296-2](https://doi.org/10.1016/0046-8177(91)90296-2)
 84. Papp MI, Kahn JE, Lantos PL (1989) Glial cytoplasmic inclusions in the CNS of patients with multiple system atrophy (striatonigral degeneration, olivopontocerebellar atrophy and Shy-Drager syndrome). *J Neurol Sci* 94:79–100. [https://doi.org/10.1016/0022-510x\(89\)90219-0](https://doi.org/10.1016/0022-510x(89)90219-0)
 85. Pastore A, Temussi PA (2012) The two faces of Janus: functional interactions and protein aggregation. *Curr Opin Struct Biol* 22:30–37. <https://doi.org/10.1016/j.sbi.2011.11.007>
 86. Paulson HL, Perez MK, Trotter Y, Trojanowski JQ, Subramony SH, Das SS, Vig P, Mandel JL, Fischbeck KH, Pittman RN (1997) Intranuclear inclusions of expanded polyglutamine protein in spinocerebellar ataxia type 3. *Neuron* 19:333–344. [https://doi.org/10.1016/s0896-6273\(00\)80943-5](https://doi.org/10.1016/s0896-6273(00)80943-5)
 87. Perbellini R, Greco S, Sarra-Ferraris G, Cardani R, Capogrossi MC, Meola G, Martelli F (2011) Dysregulation and cellular mislocalization of specific miRNAs in myotonic dystrophy type 1. *Neuromuscul Disord* 21:81–88. <https://doi.org/10.1016/j.nmd.2010.11.012>
 88. Qurashi A, Liu H, Ray L, Nelson DL, Duan R, Jin P (2012) Chemical screen reveals small molecules suppressing fragile X premutation rCGG repeat-mediated neurodegeneration in drosophila. *Hum Mol Genet* 21:2068–2075. <https://doi.org/10.1093/hmg/dds024>
 89. Rousseau F, Rouillard P, Morel ML, Khandjian EW, Morgan K (1995) Prevalence of carriers of premutation-size alleles of the FMR1 gene—and implications for the population genetics of the fragile X syndrome. *Am J Hum Genet* 57:1006–1018
 90. Saito R, Shimizu H, Miura T, Hara N, Mezaki N, Higuchi Y, Miyashita A, Kawachi I, Sanpei K, Honma Y et al (2020) Oculopharyngodistal myopathy with coexisting histology of systemic neuronal intranuclear inclusion disease: Clinicopathologic features of an autopsied patient harboring CGG repeat expansions in LRP12. *Acta Neuropathol Commun* 8:75. <https://doi.org/10.1186/s40478-020-00945-2>
 91. Salcedo-Arellano MJ, Wang JY, McLennan YA, Doan M, Cabal-Herrera AM, Jimenez S, Wolf-Ochoa MW, Sanchez D, Juarez P, Tassone F et al (2021) Cerebral microbleeds in fragile X-associated tremor/ataxia syndrome. *Mov Disord* 36:1935–1943. <https://doi.org/10.1002/mds.28559>
 92. Salcedo-Arellano MJ, Wolf-Ochoa MW, Hong T, Amina S, Tassone F, Lechpammer M, Hagerman R, Martinez-Cerdeno V (2020) Parkinsonism versus concomitant parkinson's disease in fragile X-associated tremor/ataxia syndrome. *Mov Disord Clin Pract* 7:413–418. <https://doi.org/10.1002/mdc3.12942>
 93. Satoyoshi E, Kinoshita M (1977) Oculopharyngodistal myopathy. *Arch Neurol* 34:89–92. <https://doi.org/10.1001/archneur.1977.00500140043007>
 94. Sellier C, Buijsen RAM, He F, Natla S, Jung L, Tropel P, Gaucherot A, Jacobs H, Meziane H, Vincent A et al (2017) Translation of expanded CGG repeats into FMRpolyG is pathogenic and may contribute to fragile X tremor ataxia syndrome. *Neuron* 93:331–347. <https://doi.org/10.1016/j.neuron.2016.12.016>
 95. Sellier C, Freyermuth F, Tabet R, Tran T, He F, Ruffenach F, Alunni V, Moine H, Thibault C, Page A et al (2013) Sequestration of DROSHA and DGCR8 by expanded CGG RNA repeats alters microRNA processing in fragile X-associated tremor/ataxia syndrome. *Cell Rep* 3:869–880. <https://doi.org/10.1016/j.celrep.2013.02.004>
 96. Sellier C, Rau F, Liu Y, Tassone F, Hukema RK, Gattoni R, Schneider A, Richard S, Willemsen R, Elliott DJ et al (2010) Sam68 sequestration and partial loss of function are associated with splicing alterations in FXTAS patients. *EMBO J* 29:1248–1261. <https://doi.org/10.1038/emboj.2010.21>
 97. Shi CH, Fan Y, Yang J, Yuan YP, Shen S, Liu F, Mao CY, Liu H, Zhang S, Hu ZW et al (2021) NOTCH2NLC Intermediate-length repeat expansions are associated with parkinson disease. *Ann Neurol* 89:182–187. <https://doi.org/10.1002/ana.25925>
 98. Siomi H, Siomi MC, Nussbaum RL, Dreyfuss G (1993) The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. *Cell* 74:291–298. [https://doi.org/10.1016/0092-8674\(93\)90420-u](https://doi.org/10.1016/0092-8674(93)90420-u)
 99. Sofola OA, Jin P, Qin Y, Duan R, Liu H, de Haro M, Nelson DL, Botas J (2007) RNA-binding proteins hnRNP A2/B1 and CUGBP1 suppress fragile X CGG premutation repeat-induced neurodegeneration in a drosophila model of FXTAS. *Neuron* 55:565–571. <https://doi.org/10.1016/j.neuron.2007.07.021>
 100. Sone J, Mitsuhashi S, Fujita A, Mizuguchi T, Hamanaka K, Mori K, Koike H, Hashiguchi A, Takashima H, Sugiyama H et al (2019) Long-read sequencing identifies GGC repeat expansions in NOTCH2NLC associated with neuronal intranuclear inclusion disease. *Nat Genet* 51:1215–1221. <https://doi.org/10.1038/s41588-019-0459-y>
 101. Sone J, Mori K, Inagaki T, Katsumata R, Takagi S, Yokoi S, Araki K, Kato T, Nakamura T, Koike H et al (2016) Clinicopathological features of adult-onset neuronal intranuclear inclusion disease. *Brain* 139:3170–3186. <https://doi.org/10.1093/brain/aww249>
 102. Sone J, Tanaka F, Koike H, Inukai A, Katsuno M, Yoshida M, Watanabe H, Sobue G (2011) Skin biopsy is useful for the antemortem diagnosis of neuronal intranuclear inclusion disease. *Neurology* 76:1372–1376. <https://doi.org/10.1212/WNL.0b013e3182166e13>
 103. Subramanian S, Mishra RK, Singh L (2003) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biol* 4:R13. <https://doi.org/10.1186/gb-2003-4-2-r13>
 104. Sun JH, Zhou L, Emerson DJ, Phyto SA, Titus KR, Gong W, Gilgenast TG, Beagan JA, Davidson BL, Tassone F et al (2018) Disease-associated short tandem repeats co-localize with chromatin domain boundaries. *Cell*. <https://doi.org/10.1016/j.cell.2018.08.005>
 105. Sun QY, Xu Q, Tian Y, Hu ZM, Qin LX, Yang JX, Huang W, Xue J, Li JC, Zeng S et al (2020) Expansion of GGC repeat in the human-specific NOTCH2NLC gene is associated with essential tremor. *Brain* 143:222–233. <https://doi.org/10.1093/brain/awz372>
 106. Swinnen B, Robberecht W, Van Den Bosch L (2020) RNA toxicity in non-coding repeat expansion disorders. *EMBO J* 39:e101112. <https://doi.org/10.15252/embj.2018101112>
 107. Tabrizi SJ, Leavitt BR, Landwehrmeyer GB, Wild EJ, Saft C, Barker RA, Blair NF, Craufurd D, Priller J, Rickards H (2019) Targeting huntingtin expression in patients with huntington's disease. *N Engl J Med* 380:2307–2316. <https://doi.org/10.1056/nejmoa1900907>
 108. Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH (1995) Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol* 128:995–1002. <https://doi.org/10.1083/jcb.128.6.995>
 109. Tassone F, Hagerman RJ, Taylor AK, Gane LW, Godfrey TE, Hagerman PJ (2000) Elevated levels of FMR1 mRNA in carrier males: a new

- mechanism of involvement in the fragile-X syndrome. *Am J Hum Genet* 66:6–15. <https://doi.org/10.1086/302720>
110. Tian Y, Wang JL, Huang W, Zeng S, Jiao B, Liu Z, Chen Z, Li Y, Wang Y, Min HX et al (2019) Expansion of human-specific GGC repeat in neuronal intranuclear inclusion disease-related disorders. *Am J Hum Genet* 105:166–176. <https://doi.org/10.1016/j.ajhg.2019.05.013>
 111. Todd PK, Oh SY, Krans A, He F, Sellier C, Frazer M, Renoux AJ, Chen K-c, Scaglione KM, Basrur V et al (2013) CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron* 78:440–455. <https://doi.org/10.1016/j.neuron.2013.03.026>
 112. Toko M, Ohshita T, Kurashige T, Morino H, Kume K, Yamashita H, Sobue G, Iwasaki Y, Sone J, Kawakami H et al (2021) FXTAS is difficult to differentiate from neuronal intranuclear inclusion disease through skin biopsy: a case report. *BMC Neurol* 21:396. <https://doi.org/10.1186/s12883-021-02425-z>
 113. Tomé FM, Askanas V, Engel WK, Alvarez RB, Lee CS (1989) Nuclear inclusions in innervated cultured muscle fibers from patients with oculopharyngeal muscular dystrophy. *Neurology* 39:926–932. <https://doi.org/10.1212/wnl.39.7.926>
 114. Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang FP (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 65:905–914. [https://doi.org/10.1016/0092-8674\(91\)90397-h](https://doi.org/10.1016/0092-8674(91)90397-h)
 115. Verma AK, Khan E, Mishra SK, Jain N, Kumar A (2019) Piperine modulates protein mediated toxicity in FXTAS through interacting expanded r(CG)exp RNA. *ACS Chem Neurosci* 10:3778–3788. <https://doi.org/10.1021/acscchemneuro.9b00282>
 116. Voineagu I, Surka CF, Shishkin AA, Krasilnikova MM, Mirkin SM (2009) Replisome stalling and stabilization at CGG repeats, which are responsible for chromosomal fragility. *Nat Struct Mol Biol* 16:226–228. <https://doi.org/10.1038/nsmb.1527>
 117. Wang ET, Treacy D, Eichinger K, Struck A, Estabrook J, Olafson H, Wang TT, Bhatt K, Westbrook T, Sedehizadeh S et al (2019) Transcriptome alterations in myotonic dystrophy skeletal muscle and heart. *Hum Mol Genet* 28:1312–1321. <https://doi.org/10.1093/hmg/ddy432>
 118. Wang H, Yu J, Yu M, Deng J, Zhang W, Lv H, Liu J, Shi X, Liang W, Jia Z et al (2021) GGC repeat expansion in the NOTCH2NLC gene is associated with a phenotype of predominant motor-sensory and autonomic neuropathy. *Front Genet* 12:694790. <https://doi.org/10.3389/fgene.2021.694790>
 119. Wang Y, Wang B, Wang L, Yao S, Zhao J, Zhong S, Cong L, Liu L, Zhang J, Zhang J et al (2020) Diagnostic indicators for adult-onset neuronal intranuclear inclusion disease. *Clin Neuropathol*. <https://doi.org/10.5414/NP301203>
 120. White PJ, Borts RH, Hirst MC (1999) Stability of the human fragile X (CGG)(n) triplet repeat array in *Saccharomyces cerevisiae* deficient in aspects of DNA metabolism. *Mol Cell Biol* 19:5675–5684. <https://doi.org/10.1128/mcb.19.8.5675>
 121. Woulfe JM (2007) Abnormalities of the nucleus and nuclear inclusions in neurodegenerative disease: a work in progress. *Neuropathol Appl Neurobiol* 33:2–42. <https://doi.org/10.1111/j.1365-2990.2006.00819.x>
 122. Wu C, Xiang H, Chen R, Zheng Y, Zhu M, Chen S, Yu Y, Peng Y, Yu Y, Deng J et al (2022) Genetic spectrum in a cohort of patients with distal hereditary motor neuropathy. *Ann Clin Transl Neurol* 9:633–643. <https://doi.org/10.1002/acn3.51543>
 123. Xie F, Hu X, Liu P, Zhang D (2022) A case report of neuronal intranuclear inclusion disease presenting with recurrent migraine-like attacks and cerebral edema: a mimicker of MELAS. *Front Neurol* 13:837844. <https://doi.org/10.3389/fneur.2022.837844>
 124. Xu K, Li Y, Allen E, Jin P (2021) Therapeutic development for CGG repeat expansion-associated neurodegeneration. *Front Cell Neurosci* 15:655568. <https://doi.org/10.3389/fncel.2021.655568>
 125. Yang D, Cen Z, Wang L, Chen X, Liu P, Wang H, Ouyang Z, Chen Y, Zhang F, Xie F et al (2021) Neuronal intranuclear inclusion disease tremor-dominant subtype: a mimicker of essential tremor. *Eur J Neurol*. <https://doi.org/10.1111/ene.15169>
 126. Yau WY, Vandrovцова J, Sullivan R, Chen Z, Zecchinelli A, Cilia R, Duga S, Murray M, Carmona S, Genomics England Research Consortium et al (2021) Low prevalence of NOTCH2NLC GGC repeat expansion in white patients with movement disorders. *Mov Disord* 36:251–255. <https://doi.org/10.1002/mds.28302>
 127. Yu J, Deng J, Guo X, Shan J, Luan X, Cao L, Zhao J, Yu M, Zhang W, Lv H et al (2021) The GGC repeat expansion in NOTCH2NLC is associated with oculopharyngodistal myopathy type 3. *Brain* 144:1819–1832. <https://doi.org/10.1093/brain/awab077>
 128. Yu J, Luan XH, Yu M, Zhang W, Lv H, Cao L, Meng L, Zhu M, Zhou B, Wu XR et al (2021) GGC repeat expansions in NOTCH2NLC causing a phenotype of distal motor neuropathy and myopathy. *Ann Clin Transl Neurol* 8:1330–1342. <https://doi.org/10.1002/acn3.51371>
 129. Yu J, Shan J, Yu M, Di L, Xie Z, Zhang W, Lv H, Meng L, Zheng Y, Zhao Y et al (2022) The CGG repeat expansion in RILPL1 is associated with oculopharyngodistal myopathy type 4. *Am J Hum Genet*. <https://doi.org/10.1016/j.ajhg.2022.01.012>
 130. Yuan Y, Liu Z, Hou X, Li W, Ni J, Huang L, Hu Y, Liu P, Hou X, Xue J et al (2020) Identification of GGC repeat expansion in the gene in amyotrophic lateral sclerosis. *Neurology* 95:e3394–e3405. <https://doi.org/10.1212/WNL.0000000000010945>
 131. Yudkin D, Hayward BE, Aladjem MI, Kumari D, Usdin K (2014) Chromosome fragility and the abnormal replication of the FMR1 locus in fragile X syndrome. *Hum Mol Genet* 23:2940–2952. <https://doi.org/10.1093/hmg/ddu006>
 132. Zhang N, Ashizawa T (2017) RNA toxicity and foci formation in microsatellite expansion diseases. *Curr Opin Genet Dev* 44:17–29. <https://doi.org/10.1016/j.gde.2017.01.005>
 133. Zhao X, Usdin K (2021) (Dys)function follows form: nucleic acid structure, repeat expansion, and disease pathology in FMR1 disorders. *Int J Mol Sci*. <https://doi.org/10.3390/ijms22179167>
 134. Zhong S, Lian Y, Luo W, Luo R, Wu X, Ji J, Ji Y, Ding J, Wang X (2021) Upstream open reading frame with NOTCH2NLC GGC expansion generates polyglycine aggregates and disrupts nucleocytoplasmic transport: implications for polyglycine diseases. *Acta Neuropathol* 142:1003–1023. <https://doi.org/10.1007/s00401-021-02375-3>
 135. Zhou ZD, Jankovic J, Ashizawa T, Tan EK (2022) Neurodegenerative diseases associated with non-coding CGG tandem repeat expansions. *Nat Rev Neurol*. <https://doi.org/10.1038/s41582-021-00612-7>
 136. Zu T, Gibbens B, Doty NS, Gomes-Pereira M, Huguet A, Stone MD, Margolis J, Peterson M, Markowski TW, Ingram MAC et al (2011) Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci U S A* 108:260–265. <https://doi.org/10.1073/pnas.1013343108>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

