The CGG repeat expansion in *RILPL1* is associated with oculopharyngodistal myopathy type 4

Authors

Jiaxi Yu, Jingli Shan, Meng Yu, ..., Chuanzhu Yan, Zhaoxia Wang, Jianwen Deng

Correspondence

drwangzx@163.com (Z.W.),
jianwendeng@pkufh.com (J.D.)

Yu et al., 2022, The American Journal of Human Genetics *109*, 533–541
March 3, 2022
https://doi.org/10.1016/j.ajhg.2022.01.012

CellPress

The CGG repeat expansion in *RILPL1* is associated with oculopharyngodistal myopathy type 4

Jiaxi Yu,[1,2,8] Jingli Shan,[3,8] Meng Yu,[1] Li Di,[4] Zhiying Xie,[1] Wei Zhang,[1] He Lv,[1] Lingchao Meng,[1] Yiming Zheng,[1] Yawen Zhao,[1] Qiang Gang,[1] Xueyu Guo,[5] Yang Wang,[5] Jianying Xi,[6] Wenhua Zhu,[6] Yuwei Da,[4] Daojun Hong,[7] Yun Yuan,[1] Chuanzhu Yan,[3,9] Zhaoxia Wang,[1,2,9,*] and Jianwen Deng[1,2,9,*]

Summary

Recent studies indicate that CGG repeat expansions in *LRP12*, *GIPC1*, and *NOTCH2NLC* are associated with oculopharyngodistal myopathy (OPDM) types 1, 2, and 3, respectively. However, some clinicopathologically confirmed OPDM cases continue to have unknown genetic causes. Here, through a combination of long-read whole-genome sequencing (LRS), repeat-primed polymerase chain reaction (RP-PCR), and fluorescence amplicon length analysis PCR (AL-PCR), we found that a CGG repeat expansion in the 5′ UTR of *RILPL1* is associated with familial and simplex OPDM type 4 (OPDM4). The number of repeats ranged from 139 to 197. Methylation analysis indicates that the methylation levels in *RILPL1* were unaltered in OPDM4 individuals. Analyses of muscle biopsies suggested that the expanded CGG repeat might be translated into a toxic poly-glycine protein that co-localizes with p62 in intranuclear inclusions. Moreover, analyses suggest that the toxic RNA gain-of-function effects also contributed to the pathogenesis of this disease. Intriguingly, all four types of OPDM have been found to be associated with the CGG repeat expansions located in 5′ UTRs. This finding suggests that a common pathogenic mechanism, driven by the CGG repeat expansion, might underlie all cases of OPDM.

Oculopharyngodistal myopathy (OPDM) is a rare hereditary neuromuscular disease that was first described in 1977.[1] OPDM-affected individuals present with progressive ptosis, ophthalmoparesis, facial and masseter weakness, dysphagia, and muscle weakness of distal limbs. The major myopathological features showed chronic myopathic changes with rimmed vacuoles (RVs) and filamentous intranuclear inclusions.[1–3] The diagnosis is currently based on typical clinical features, myopathological findings, and exclusion of other neuromuscular disorders with overlapping phenotypes by genetic analysis.

Recent studies have indicated that OPDM exhibits marked genetic heterogeneity. It can be caused by a CGG repeat expansion in the 5′ untranslated region (5′ UTR) of three genes: LDL receptor-related protein 12 (*LRP12*) (MIM: 618299),[4] GAIP/RGS19-interacting protein (*GIPC1*) (MIM: 605072),[5,6] and notch 2 N-terminal-like C (*NOTCH2NLC*) (MIM: 618025),[7,8] associated with OPDM1 (MIM: 164310), OPDM2 (MIM: 618940), and OPDM3 (MIM: 619473), respectively. Among these genes, CGG repeat expansion in *NOTCH2NLC* was also found to be associated with neuronal intranuclear inclusion disease (NIID, OMIM: 603472),[4,9–11] suggesting a NIID-OPDM spectrum disorder that most likely shares similar pathogenic mechanisms. However, there remains no known genetic cause for disease in a proportion of OPDM-affected individuals. We hypothesized that these individuals might also have a CGG repeat expansion in the 5′ UTR of yet unknown gene(s) on the basis of the fact that all three types of genetically known OPDM were associated with the expansion of CGG repeats in the 5′ UTR.[4–8]

In this study, we collected the clinical and pathological data of two unrelated OPDM-affected individuals (F1-III-1, affected by familial OPDM, and S1, affected by a simplex occurrence) (Figure 1A) with unknown genetic causes of disease. The procedures followed were approved by the Local Ethics Committee of Peking University First Hospital, and the proper informed consents were obtained.

The index individual (F1-III-1) from family 1 is a 43-year-old female who presented bilateral ptosis, limitation of eye movement, and diplopia at the age of 26, then gradually developed facial muscle weakness, a nasal voice, difficulty in speaking and swallowing, and symmetrical weakness in four distal limbs. Physical examination at the age of 43 revealed postural tremor of both hands, in addition to muscle weakness and atrophy in facial muscles, both hands, and lower limbs. Deep tendon reflexes were absent. Light touch sensation, pinprick sensation, and proprioception were intact. Serum CK level was 268 IU/L (normal range: 24–194 IU/L). Electromyography revealed features of myogenic changes. Muscle MRI

[1]Department of Neurology, Peking University First Hospital, Beijing, 100034, China; [2]Beijing Key Laboratory of Neurovascular Disease Discovery, Beijing, 100034, China; [3]Research Institute of Neuromuscular and Neurodegenerative Diseases and Department of Neurology, Qilu Hospital, Shandong University, Jinan, 250012, China, Mitochondrial Medicine Laboratory, Qilu Hospital (Qingdao), Shandong University, Qingdao, China, Brain Science Research Institute, Shandong University, Jinan, China; [4]Department of Neurology, Xuanwu Hospital, Capital Medical University, Beijing, 100053, China; [5]Grandomics Biosciences, Beijing, 100000, China; [6]Department of Neurology, Huashan Hospital Fudan University, Shanghai, 200040, China; [7]Department of Neurology, The First Affiliated Hospital of Nanchang University, Nanchang, 330006, China
[8]These authors contributed equally to this work and are joint first authors
[9]These authors contributed equally to this work and are joint senior authors
*Correspondence: drwangzx@163.com (Z.W.), jianwendeng@pkufh.com (J.D.)
https://doi.org/10.1016/j.ajhg.2022.01.012.

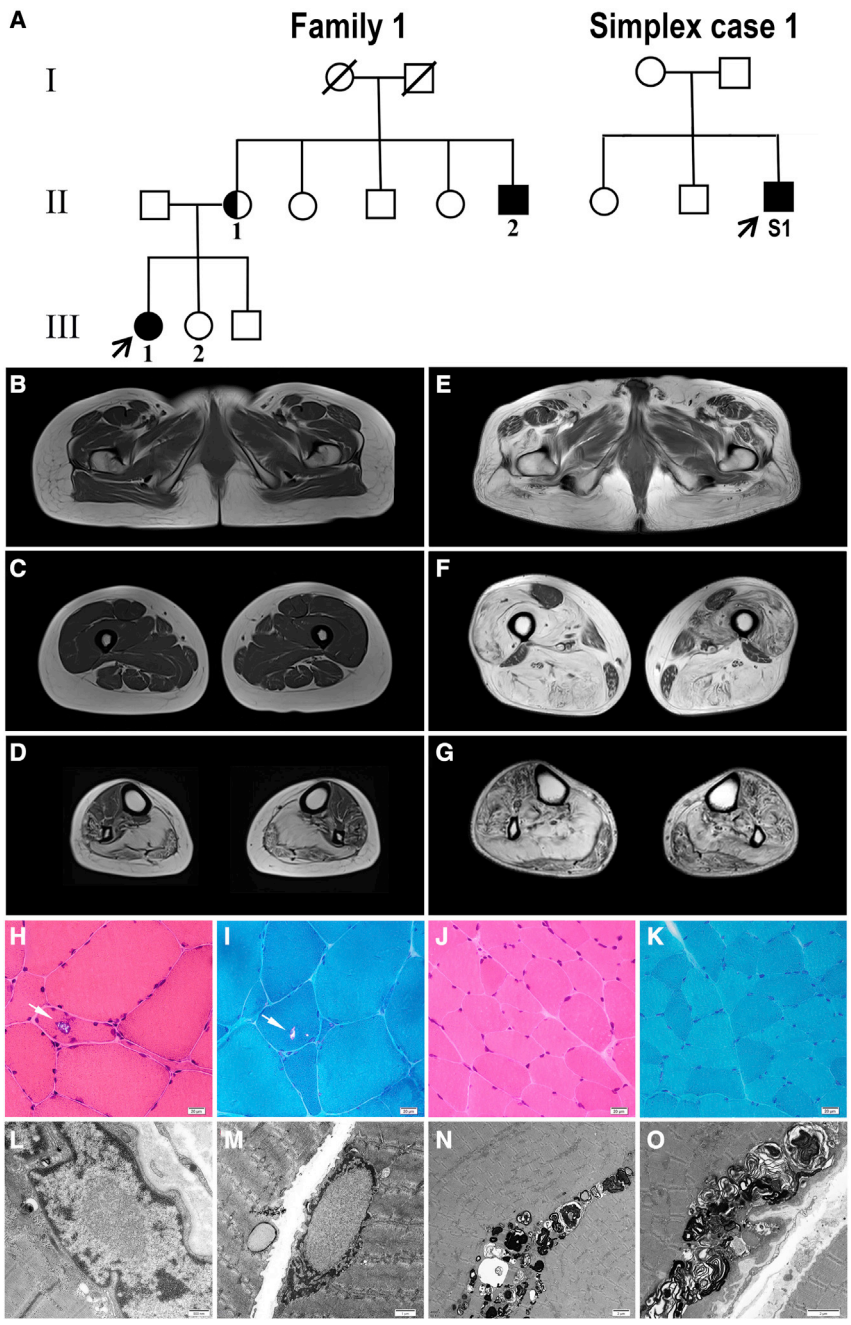

Figure 1. Muscle MRI and myopathological changes of the OPDM4-affected individuals

(A) Pedigrees of family 1 and OPDM-affected simplex individual 1. OPDM-affected individuals are indicated by filled symbols. The individual with essential tremor but not OPDM is indicated by a half-filled symbol. All numbered individuals had available blood DNA.

(B–G) Different extents of fatty infiltration of lower-limb muscles by muscle MRI in F1-III-1 (B–D) and S1 (E–G).

(H–K) Variation in fiber size and rimmed vacuoles (marked by an arrow) on muscle biopsy in S1 (H and I) compared with the control (J and K).

(L–O) Intranuclear filamentous inclusions (L and M) and various myelin figures and autophagic vacuoles (N and O) of muscle tissue upon electron microscopy in F1-II-2 (L) and S1 (M–O).

(H and J) H&E staining; (I and K) mGT staining. Scale bars represent 20 μm (H–K), 500 nm (L), 1 μm (M), and 2 μm (N and O).

performed at age 43 showed that the pelvic and thigh muscles were almost intact, whereas the soleus and the gastrocnemius muscles had severe fat infiltration (Figures 1B–1D).

In family 1, one of the proband's maternal uncles, F1-II-2, displayed the same symptoms. F1-II-2 had exhibited ptosis since the age of 27. Examination at age 42 revealed he had poor eye movement, a nasal voice, and muscle weakness in facial muscles and distal limbs. No sensory impairment was noted. Electromyography showed myogenic changes. The proband's mother, F1-II-1, did not show any sign of muscle weakness except the presence of a postural tremor in her head and both hands.

The simplex individual 1 (S1) is a 46-year-old male who noticed drooping eyelids at the age of 14, then gradually developed limitation of eye movement. At age 31, he noticed weakness in finger extension. After the age of 36, he gradually developed dysarthria and weakness of distal lower limbs. Family history was denied. Physical examination at age 46 revealed a high-arched palate, external ophthalmoplegia, facial weakness, and wasting in both lower limbs and distal upper limbs. Deep tendon reflexes were absent. No sensory deficit was found. Muscle MRI performed at age 46 showed severe fatty replacements in both thigh and calf muscles, with rectus femoris, sartorius, gracilis, short head of the biceps femoris, and gastrocnemius muscle relatively spared (Figures 1E–1G). Serum CK was 477 IU/L (normal range: 70–170 IU/L). Electromyography revealed myogenic changes.

Muscle biopsies of three OPDM-affected individuals (F1-II-2, F1-III-1, and S1) showed variation in fiber size and rimmed vacuoles on both Haematoxylin and eosin (H&E) staining and modified Gomori trichrome (mGT) staining in comparison to the control (Figures 1H–K and Figure S1). Electron microscopy of muscle samples revealed intranuclear filamentous inclusions without limiting membrane (Figures 1L and 1M), as well as various myelin figures and autophagic vacuoles containing osmiophilic deposits and amorphous materials (Figures 1N and 1O).

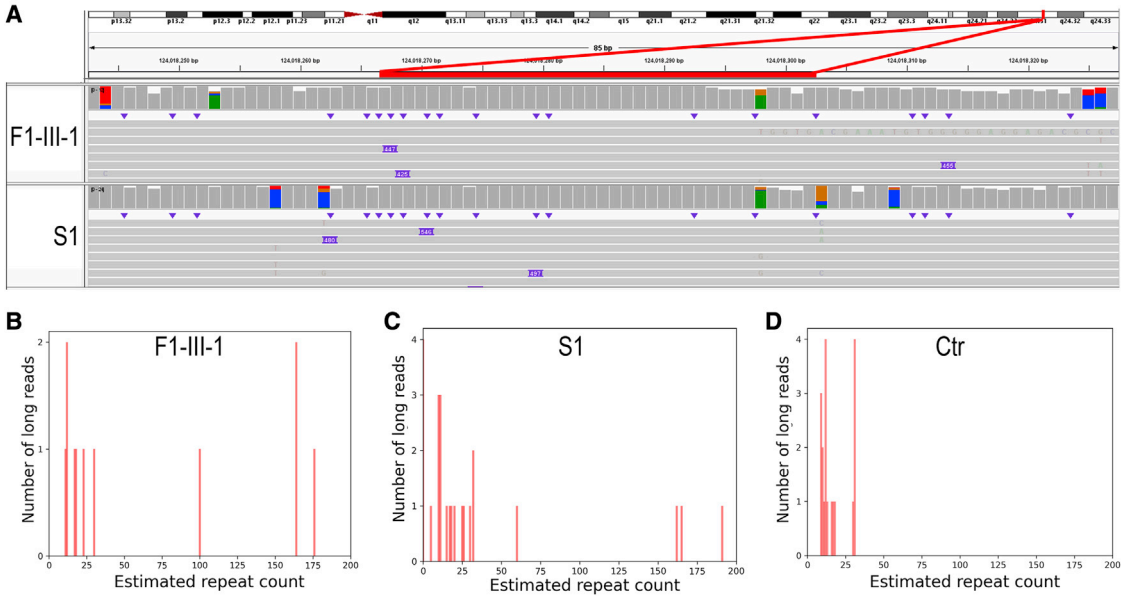

Figure 2. Identification of the CGG repeat expansion and repeat sizes in *RILPL1* by LRS in OPDM-affected individuals
(A)The Integrative Genomics Viewer showed that three lanes carried the *RILPL1* CGG repeat expansion (chr12: 124,018,267–124,018,302, hg19 version), as opposed to the normal allele, in in F1-III-1 and S1.
(B–D) LRS data showed an estimated CGG repeat count of *RILPL1* of more than 100 in OPDM-affected individuals F1-III-1 (B) and S1 (C) but no more than 40 in the controls (D).

Next, we performed long-read whole-genome sequencing (LRS) by using a PromethION sequencer (Oxford Nanopore Technologies) on the two unrelated OPDM-affected individuals (F1-III-1 and S1) for whom the known pathogenic coding variants and disease-causing repeat expansions were excluded by whole-exome sequencing (WES) and genetic analysis, respectively. We analyzed the sequencing data via our previous methods.[5,7] We searched for candidate genes with CGG repeats in their 5′ UTRs and identified a heterozygous CGG repeat expansion that was upstream of *RILPL1* (MIM: 614092) and had was more than 100 repeats long in both F1-III-1 and S1 (Figure 2 and Figure S2). Re-analysis of our previous skeletal muscle RNA-Seq data from normal controls[5] indicated that the CGG repeat expansion could be located in the 5′ UTR region of *RILPL1* (Figure S3). The transcription start site of this transcript might be located further upstream of the currently reported transcript form (GenBank: NM_178314.5, from RefSeq). *RILPL1* encodes a Rab interacting lysosomal protein-like 1,[12] which was reported to participate in the regulation of ciliary membrane content by promoting protein removal from the primary cilium.[13] Recent studies indicated that RILPL1 plays an important role in the pathogenesis of *LRRK2* mutations by interacting with RAB10.[14,15] To our knowledge, mutations of *RILPL1* have not previously been linked to any genetic disorder.

We performed repeat-primed polymerase chain reaction (RP-PCR) as previously reported[9] to confirm the LRS results and identify the CGG repeat expansion in *RILPL1* in additional individuals with OPDM (supplemental material and methods). A saw-tooth pattern demonstrated by the electropherogram was observed in three individuals exhibiting familial OPDM (F1-II-1, F1-II-2, and F1-III-1) and in a simplex individual (S1), but not in the unaffected individual in family 1 (F1-III-2) (Figure 3A-B). This analysis confirmed the CGG repeat expansion mutation of *RILPL1* in these affected individuals and indicated a familial co-segregation of the *RILPL1* variant in family 1. Then we performed RP-PCR analysis on 20 other OPDM-affected individuals with unknown genetic causes and 200 unaffected Chinese controls. CGG repeat expansion upstream of *RILPL1* was identified in an additional nine OPDM-affected individuals, whereas none of the 200 normal Chinese controls showed repeat expansion. Altogether, 11 out of 22 unrelated OPDM-affected individuals (CGG repeat expansions in *LRP12*, *GIPC1*, and *NOTCH2NLC* were excluded) carried a CGG repeat expansion in *RILPL1*.

Subsequently, we determined the CGG repeat size of *RILPL1* by fluorescence amplicon-length-analysis PCR (AL-PCR) (Figure 3C, supplemental material and methods in Document S1). The result showed that the size of the CGG repeat ranged from 139 to 197 (169.91 ± 21.82) repeats in individuals with OPDM, and 9 to 16 repeats in 200 control individuals (Figure 3D). Notably, the number of repeats in F1-II-1 was 86, which is an intermediate size between normal control and OPDM-affected individuals.

Together, these results strongly indicated that CGG repeat expansion in the 5′ UTR of *RILPL1* was associated with OPDM type 4 (OPDM4). Additionally, through the study of 11 OPDM4-affected individuals, we determined that there was no correlation between the number of repeats and the age at onset (Figure 3E).

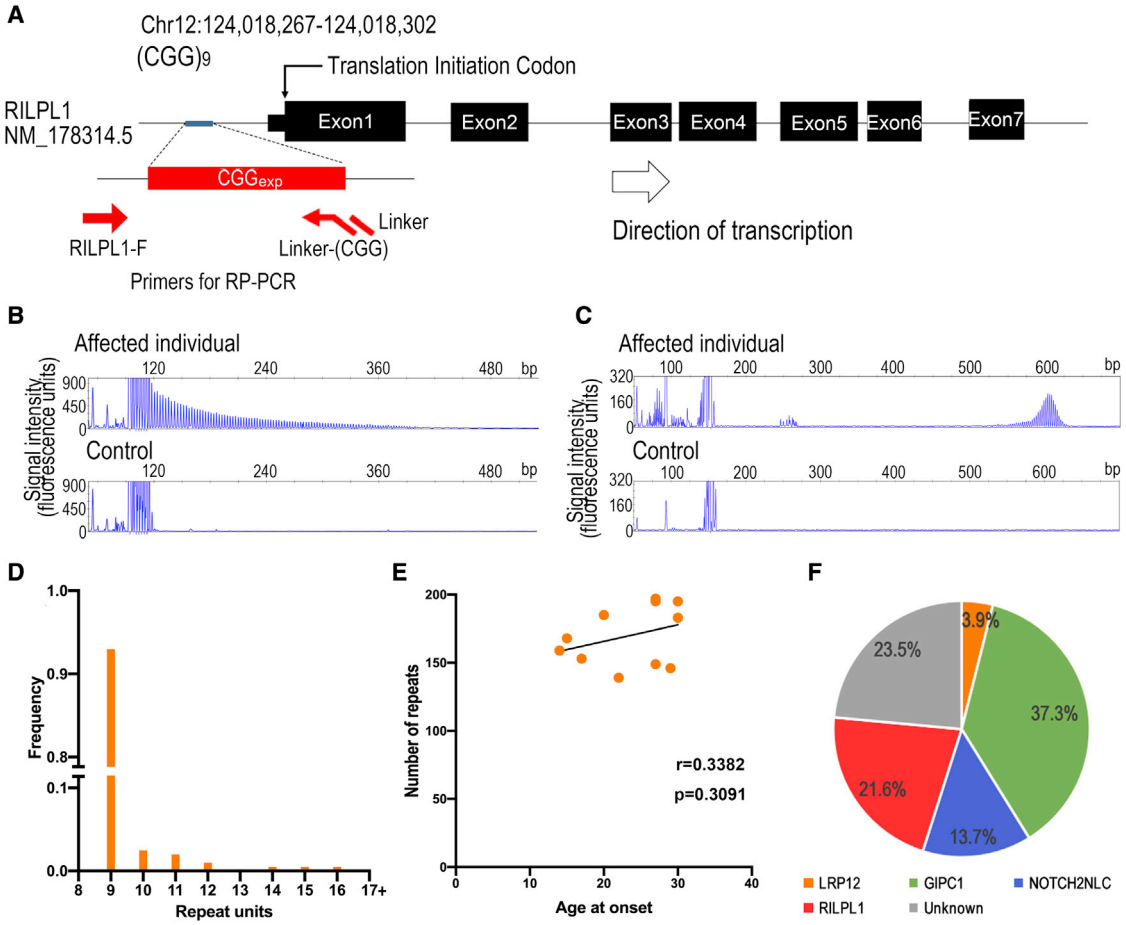

Figure 3. Validation of the CGG repeat expansion and repeat sizes in *RILPL1* by RP-PCR and AL-PCR in OPDM-affected individuals
(A) Schematic representation of *RILPL1*. The disease-associated repeat expansion (red boxes) was identified in the 5′ UTR. The primer set was designed for RP-PCR analysis to identify the CGG repeat expansion (red lines and arrows).
(B) Representative results of RP-PCR analysis showing CGG repeat expansion in *RILPL1* in S1. No CGG repeat expansion was found in the unaffected control. Experiments were conducted three times with reproducible results.
(C) Representative results of AL-PCR analysis showing the repeat count of expanded CGG in *RILPL1* in S1. The number of repeats was multiplied by 3 bp per repeat. No CGG repeat expansion was detected in the unaffected control.
(D) Fragment analysis showed that the frequency distribution of CGG repeat units in *RILPL1* ranged from 9 to 16 among 200 normal controls.
(E) There was no significant correlation between the number of CGG repeats and the age at onset in 11 Chinese OPDM4-affected individuals ($r = 0.3382$, $p = 0.3091$, $n = 11$).
(F) Pie chart for the percentages of disease-causing gene mutations and unknown gene mutations in three cohorts of Chinese OPDM-affected individuals. Trinucleotide repeat expansions in the 5′ UTRs of *LRP12*, *GIPC1*, *NOTCH2NLC*, and *RILPL1* accounted for 3.9%, 37.3%, 13.7%, and 21.6% of disease cases in 51 unrelated Chinese OPDM-affected individuals, respectively, whereas unknown genetic causes accounted for 23.5% of disease cases in this cohort.

In total, we recruited 51 unrelated Chinese OPDM-affected individuals from three neuromuscular diseases centers to explore their genotype distribution. Trinucleotide repeat expansions in *LRP12*, *GIPC1*, *NOTCH2NLC*, and *RILPL1* accounted for 3.9%, 37.3%, 13.7%, and 21.6%, respectively, leaving the remaining 23.5% with a yet undetermined genetic cause of disease (Figure 3F). A comparison of the clinical features and pathological changes among OPDM1, OPDM2, OPDM3, and OPDM4 is presented in Table 1. In brief, there is no significant difference in age of onset, disease duration, onset age of distal-limb weakness, serum CK level, or percentage of RVs among four subgroups of OPDM. However, 90% of OPDM4 affected individuals presented with ptosis or dysarthria as their onset symptoms, and 60% of OPDM4-affected individuals didn't develop distal-limb weakness until 10 to 20 years after disease onset. On the contrary, distal-limb weakness was the most common initial symptom in OPDM1 (75%), OPDM2 (73.68%), and OPDM3 (87.5%) (Table 1). These data indicate that the distal-limb muscle weakness observed in OPDM4 developed more slowly than in the other three subtypes.

To investigate the interruptions in CGG repeat expansions, we performed LRS on two OPDM4-affected individuals (simplex individuals S2 and S3). In addition to CGG repeat expansions, two forms of repeat unit including AGG and TGG were observed (Figure S2A). The CGG repeats were predominant in the repeat expansions of four

Table 1. Comparison of clinical and pathological features between OPDM1, OPDM2, OPDM3, and OPDM4

Variables	OPDM1 (n = 2)	OPDM2 (n = 19)	OPDM3 (n = 7)	OPDM4 (n = 11)	p value
Number of OPDM-affected individuals	4	27	8	11	-
Sex (male/female)	3/1	18/9	3/5	7/4	-
Age of onset (years)	32.75 ± 4.57	29.08 ± 10.33	23.13 ± 5.96	23.80 ± 6.23	ns
Disease duration (years)	10.50 ± 8.54	12.48 ± 6.84	14.57 ± 5.26	19.30 ± 7.45	ns
Clinical manifestations					
Ptosis	3/4 (75%)	21/25 (84%)	8/8 (100%)	10/10 (100%)	-
External ophthalmoplegia	3/4 (75%)	17/25 (68%)	5/8 (62.5%)	8/10 (80%)	-
Dysphagia	3/4 (75%)	18/25 (72%)	7/8 (87.5%)	8/10 (80%)	-
Dysarthria	3/4 (75%)	24/25 (96%)	8/8 (100%)	9/10 (90%)	-
Facial muscle weakness	3/4 (75%)	24/25 (96%)	8/8 (100%)	9/10 (90%)	-
Distal limb weakness	4/4 (100%)	25/25 (100%)	8/8 (100%)	10/10 (100%)	-
Onset with distal limb weakness	3/4 (75%)	14/19 (73.68%)	7/8 (87.5%)	1/10 (10%)	<0.05
Onset age of distal limb weakness (years)	33.25 ± 4.86	31.40 ± 9.95	24.00 ± 7.58	35.20 ± 9.08	ns
The duration from disease onset to the development of distal limb weakness (years)	0.50 ± 1.00	1.55 ± 3.24	1.00 ± 2.83	11.90 ± 6.14	<0.0001
Tremor	0/4 (0%)	0/25 (0%)	0/5 (0%)	1/10 (10%)	-
Serum CK[a] (IU/L) (number of available individuals)	669.75 ± 418.36 (4)	843.02 ± 591.72 (17)	638.86 ± 465.99 (7)	375.20 ± 191.94 (10)	ns
EMG pattern (number of available individuals)	MC (4)	MC (15); MC + NC (1)	MC (5); NC (2); MC + NC (1)	MC (10)	-
Rimmed vacuoles (RVs)	4/4 (100%)	19/19 (100%)	8/8 (100%)	10/10 (100%)	-
Percentage of RVs (%)	2.64% ± 2.17%	3.83% ± 2.59%	4.65% ± 3.88%	1.74% ± 1.73%	ns
Genetic cause	LRP12	GIPC1	NOTCH2NLC	RILPL1	-
Repeat-expansion size (number of available individuals)	115.00 ± 12.73 (2)	113.50 ± 19.44 (24)	169.50 ± 30.88 (4)	167.64 ± 25.26 (11)	-

Abbreviations are as follows: OPDM1/2/3/4 = oculopharyngodistal myopathy type1/type2/type3/type4; M = male; F = female; NA = not available; EMG = electromyogram; MC = myopathic change; and NC = neurogenic change.
p values are shown as OPDM4 versus OPDM1, OPDM2, and OPDM3.
ns = no significance.
[a]Normal limits: 70–170 IU/L.

OPDM4-affected individuals (F1-III-1, S1, S2, and S3), whereas the frequency of AGG and TGG interruptions was low (Figure S2B).

CGG repeat expansion in the 5′ UTR frequently leads to hypermethylation and downregulates gene expression.[16,17] Here, we investigated the 5-methylcytosine (5mC) modification of the CGG repeats and the adjacent CpG islands of *RILPL1* by using the ONT PromethION sequencing data from four OPDM4-affected individuals (F1-III-1, S1, S2, and S3). The methylation level around the CGG repeats was not high. No significant differences were detected between OPDM4-affected individuals and healthy individuals ($p = 0.4031$, Figures 4A and 4B). Subsequently, 5mC modifications of the expanded and non-expanded alleles were analyzed, and no significant difference was detected ($p = 0.0594$, Figures 4C and 4D). Consistently, the RILPL1 abundance was not significantly different between OPDM4-affected individuals and age-matched controls (Figures 4E and 4F). Given that OPDM is a progressive neuromuscular disease, we examined the abundance of RILPL1 at different ages. The results indicated that there was no change in the abundance of RILPL1 in skeletal muscle during development and aging, as shown in three different age groups (Figures 4G and 4H). Thus, the toxic effects of CGG repeat expansion might not be related to the RILPL1 abundance.

Subsequently, we carried out immunofluorescence analysis on muscle sections from F1-II-2 and S1 and an age-matched control with anti-p62 (Abcam, ab56416) and anti-RILPL1 (Proteintech, 16732-1-AP) antibodies to examine the distribution of RILPL1 in muscle samples. RILPL1 was distributed diffusely in the skeletal muscle from both control and OPDM4-affected individuals. No obvious co-localization of RILPL1 and p62 was observed in either intranuclear inclusions or RVs (Figure S4). This result suggests that the pathogenic effect of CGG repeat expansion might be independent of RILPL1.

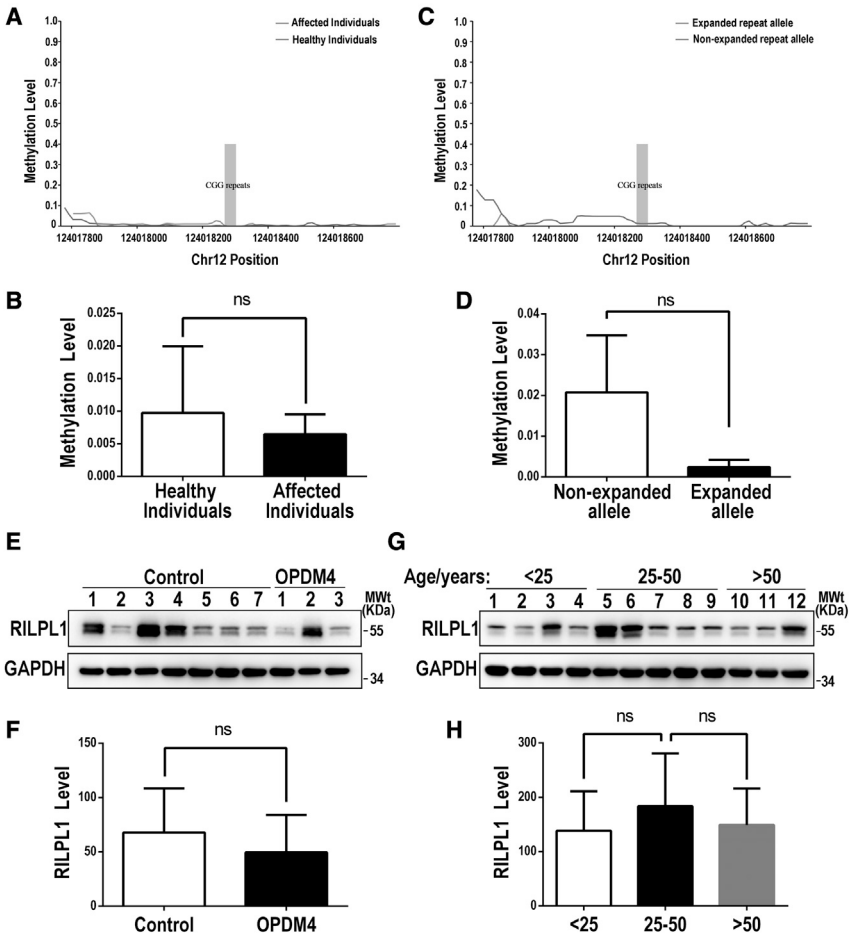

Figure 4. Methylation status analysis and immunoblot on muscle biopsy samples

(A) Methylation status across the expanded CGG repeat region in *RILPL1* from whole-blood DNA was determined with LRS data from four OPDM4-affected individuals (F1-III-1, S1, S2, and S3) and 10 healthy individuals. No significant difference in methylation was detected between affected individuals and healthy individuals.

(B) Quantification of methylation levels between OPDM4-affected individuals and healthy individuals.

(C) Methylation status between expanded alleles and non-expanded alleles in four OPDM4-affected individuals (F1-III-1, S1, S2 and S3). No significant difference in methylation was detected between expanded and non-expanded alleles.

(D) Quantification of methylation levels between expanded alleles and non-expanded alleles.

(E) Immunoblot analysis of RILPL1 in three OPDM4-affected individuals with expanded CGG repeats in *RILPL1* and seven age-matched control subjects. GAPDH was used as a loading control. Experiments were conducted thrice with reproducible results.

(F) Quantification of relative RILPL1 abundance in each group; no significant difference was observed between OPDM4 affected individuals and controls. (G) Immunoblot analysis of RILPL1 in 12 control subjects at different ages; these included four control subjects under 25 years old, five control subjects 25 to 50 years old, and three control subjects over 50 years old. GAPDH was used as a loading control. Experiments were conducted thrice with reproducible results.

(H) Quantification of relative RILPL1 abundance in each group; no significant difference was observed among control individuals at different ages. Data were analyzed with a Student's t test; ns, not significant.

It has been reported that the CGG repeat expansions in the 5′ UTR of *FMR1* (MIM: 309550) and *NOTCH2NLC* are translated into a novel poly-glycine (polyG) protein,[18,19] representing polyG disease. Our previous study indicated that an anti-glycine (Abcam ab9442) antibody is able to recognize polyG protein in muscle samples from individuals with *NOTCH2NLC*-related OPDM3.[7] To investigate whether the CGG repeat expansion in *RILPL1* could trigger a translation of polyG protein, immunofluorescence was carried out with anti-p62 (Abcam, ab56416) and anti-glycine (Abcam ab9442) antibodies. We observed co-localization of glycine and p62 in the intranuclear inclusions of skeletal-muscle samples from OPDM4-affected individuals but not in the age-matched control individual (Figure 5A). Consistent with our previous study,[7,19] polyG tends to form aggregates in the nucleus and co-localizes with p62. Thus, our data suggest that the CGG repeats in *RILPL1* might be translated into polyG protein and deposited in the nucleus. Interestingly, glycine staining signals were significantly increased in muscle fibers of OPDM4-affected individuals compared with the control individual (Figure S5), which further supported the hypothesis that expanded CGG repeats in this gene might lead to the production of polyG and its accumulation in muscle. These data suggest that the CGG repeats might be translated into polyG protein and deposited in the nucleus, which might play a role in the molecular pathogenesis of OPDM4.

Then, we designed two specific probes complementary to the upstream or downstream sequences adjacent to the expanded repeats in *RILPL1* mRNA (Figure 5B and supplemental material and methods) and performed RNA fluorescence *in situ* hybridization (FISH) combined with immunofluorescence to explore possible RNA toxicity on muscle samples from control and OPDM4-affected individuals. We observed that *RILPL1* mRNA with expanded repeats formed RNA foci, sequestered RNA binding protein muscleblind-like 1 (MBNL1), and co-localized with p62 in the intranuclear inclusions in OPDM4-affected individuals but not in the control individual (Figure 5C).

Our analyses revealed CGG repeat sizes ranging from 139 to 197 repeats (median at 168) in individuals with OPDM4. This repeat size is similar to that seen in *NOTCH2NLC*-related OPDM3,[7,8] in which most of the pathogenic CGG repeat sizes were close to 200 repeats. Meanwhile, the CGG repeat sizes in *NOTCH2NLC* in the

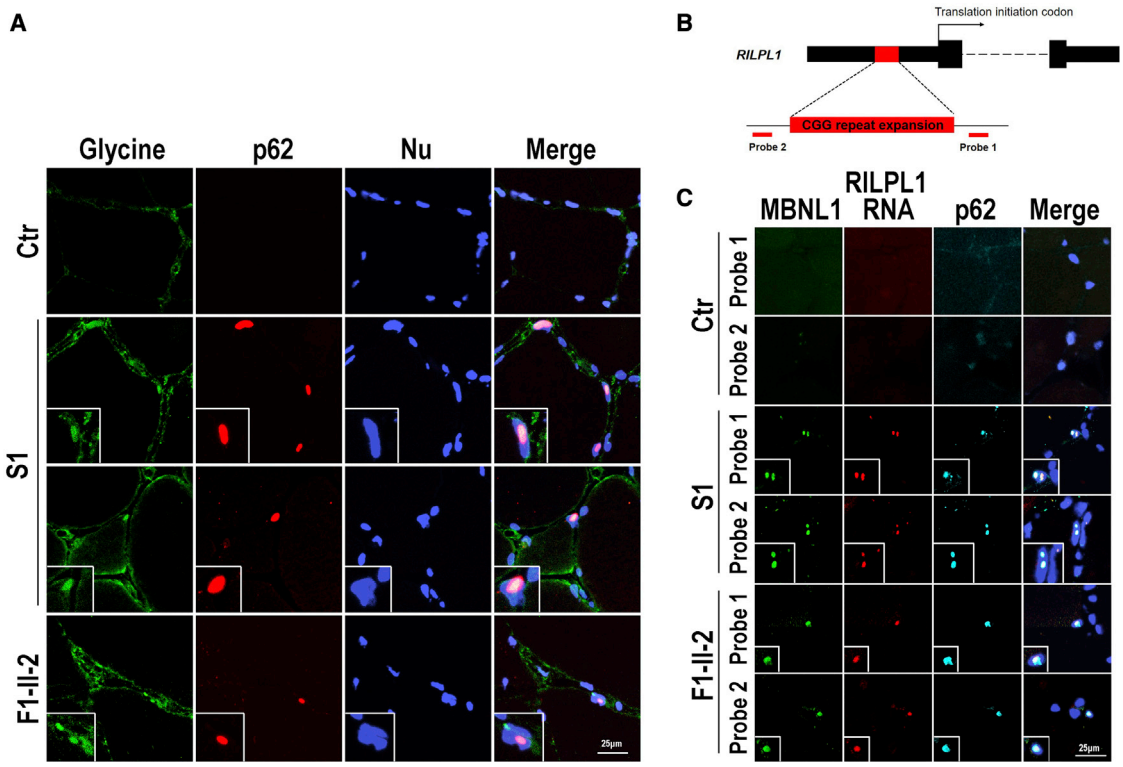

Figure 5. Immunofluorescence and RNA FISH on muscle biopsy samples
(A) Immunofluorescence showing glycine and p62 co-localized in the intranuclear inclusions of OPDM4-affected individuals (F1-II-2 and S1) but not in the control individual (higher-magnification images of representative co-localization were shown in the insets of the corresponding panels). The scale bar represents 25 μm.
(B) Probes 1 and 2 were complementary to the adjacent downstream and upstream regions of repeat expansions in *RILPL1* mRNA.
(C) RNA FISH combined with immunofluorescence indicated that expanded repeats in *RILPL1* mRNA formed RNA foci (labeled by probe 1 or probe 2) and were co-localized with MBNL1 and p62 in the intranuclear inclusions of OPDM4-affected individuals (F1-II-2 and S1) but not in the control individual (higher-magnification images of representative co-localization are shown in the insets of the corresponding panels). The scale bar represents 25 μm.

essential tremor (MIM: 618866) group were significantly less than those of the NIID group.[20] Coincidently, in this study, F1-III-1's mother (F1-II-1), who presented with essential tremor but not OPDM, also carried a shorter length of CGG repeat expansion in *RILPL1* (86 repeats). The milder phenotype with shorter repeat expansion in F1-II-1 might be due to incomplete penetrance. It also suggested that different lengths of CGG repeat expansion in *RILPL1* might be associated with variable clinical phenotypes, as observed in the NOTCH2NLC-related repeat expansion disorder.[21] Here we did not observe the correlation between CGG repeat size in *RILPL1* and onset age. Because the number of our OPDM-affected individuals was limited, it is necessary to enroll more OPDM-affected individuals in future studies.

Given that the four OPDM genotypes showed similar clinical, pathological, and genetic features, they might share a common pathogenic mechanism. CGG repeat expansion in the 5′ UTR can cause hypermethylation, which could play a vital role in pathogenesis in some diseases.[17] However, our data indicate that the methylation level around CGG repeats did not significantly differ between the expanded allele and the non-expanded allele of *RILPL1*. This result is consistent with results observed for OPDM1, OPDM2, and OPDM3.[4,5,7] Conversely, the CGG repeat expansion in *NOTCH2NLC* leads to high methylation levels and decreases gene expression in asymptomatic carriers with more than 300 CGG repeats.[7,22] So it could be deduced that CGG repeat expansion in the 5′ UTR does not cause disease through hypermethylation in OPDM.

Toxic protein gain-of-function mediated by repeat-associated non-AUG (RAN) translation or canonical AUG translation is also a potential molecular mechanism underlying OPDM.[17] Recent studies showed that for *NOTCH2NLC*-related OPDM3 and NIID, CGG repeat expansion was translated into a pathogenic polyGlycine-containing protein that was toxic in cellular and animal models.[7,19] A recent review of OPDM suggested that the expanded CGG repeat in the 5′ UTR might be translated into toxic polyG proteins.[17] Here, our data indicated that expanded CGG repeats in the 5′ UTR of *RILPL1* might be translated into polyG protein as well. This suggests that OPDM4 might have a toxic protein gain-of-function mechanism, representing a distinct class of disorder, polyG disease. However, there are limitations to our current study: (1) whether the *RILPL1* CGG repeats are embedded in an

upstream open-reading frame (uORF) that encodes a harmful polyG containing peptide remains to be investigated; and (2) the molecular mechanisms by which RILPL1-polyG causes degeneration of muscle cells is still unclear. Thus, additional studies on cellular and animal models are required for validation of RILPL1-polyG production and its pathogenic role in OPDM.

Toxic RNA gain-of-function was also reported to be involved in the pathogenic mechanism of OPDM3.[7,22,23] Recent studies showed that CGG-expanded RNA foci sequestered multiple RNA binding proteins to form insoluble inclusions in the nucleus.[7] In this study, we found that a repeat expansion in *RILPL1* mRNA formed RNA foci and sequestered RNA binding protein MBNL1 into intranuclear inclusions. MBNL1 was also reported to be sequestered in the intranuclear inclusions and induced RNA toxicity in myotonic dystrophy 1 (DM1, MIM: 160900), NIID, and fragile X tremor/ataxia syndrome (FXTAS, MIM: 300623), indicating that the toxic RNA gain-of-function mechanism also played an important role in the pathogenesis of OPDM4.[18,22,24]

In conclusion, we identified the CGG repeat expansion in the 5′ UTR of *RILPL1* as the causative mutation of OPDM4. This finding will provide further understanding into the molecular basis of OPDM, which might be caused by a similar molecular pathway with different genotypes.

Data and code availability

The datasets supporting this study have not been deposited in a public repository because of ethical restriction. Original data are available from the corresponding author upon reasonable request.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2022.01.012.

Acknowledgments

We are indebted to the cooperation of the OPDM-affected individuals and their families. We thank Mr. Lijun Chai and Mr. Jin Xu (Peking University First Hospital) for their work in taking electron microscopy pictures and Ms. Jing Liu and Ms. Qingqing Wang (Peking University First Hospital) for their work in preparations for pathological sections. The work was funded by the National Natural Science Foundation of China (grants 81571219, 82071409, 82171846, 82160252, and U20A20356), the Double Thousand Talents Program of Jiangxi Province (jxsq2019101021), the Science and Technology Project of the Jiangxi Health Committee (202110028), the Peking University Medicine Fund of Fostering Young Scholars' Scientific and Technological Innovation (BMU2021PY003), and the Innovation Fund for Outstanding Doctoral Candidates of Peking University Health Science Center (to J.Y.).

Declaration of interests

The authors declare no competing interests.

Received: October 12, 2021
Accepted: January 20, 2022
Published: February 10, 2022

Web resources

GenBank, https://www.ncbi.nlm.nih.gov/genbank/
Integrative Genomics Viewer, https://software.broadinstitute.org/software/igv/
OMIM, https://omim.org
RepeatHMM, https://github.com/WGLab/RepeatHMM
UCSC Genome Browser, https://genome.ucsc.edu

References

1. Satoyoshi, E., and Kinoshita, M. (1977). Oculopharyngodistal myopathy. Arch. Neurol. *34*, 89–92. https://doi.org/10.1001/archneur.1977.00500140043007%J.
2. Durmus, H., Laval, S.H., Deymeer, F., Parman, Y., Kiyan, E., Gokyigiti, 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.
3. Lu, H., Luan, X., Yuan, Y., Dong, M., Sun, W., and Yan, C. (2008). The clinical and myopathological features of oculopharyngodistal myopathy in a Chinese family. Neuropathology *28*, 599–603. https://doi.org/10.1111/j.1440-1789.2008.00924.x.
4. Ishiura, H., Shibata, S., Yoshimura, J., Suzuki, Y., Qu, W., Doi, K., Almansour, M.A., Kikuchi, J.K., 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.
5. 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.
6. Xi, J., Wang, X., Yue, D., Dou, T., Wu, Q., Lu, J., Liu, Y., Yu, W., Qiao, K., Lin, J., et al. (2021). 5′ UTR CGG repeat expansion in GIPC1 is associated with oculopharyngodistal myopathy. Brain *144*, 601–614. https://doi.org/10.1093/brain/awaa426.
7. 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.
8. 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.
9. 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.
10. 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.

11. Tian, Y., Wang, J.L., Huang, W., Zeng, S., Jiao, B., Liu, Z., Chen, Z., Li, Y., Wang, Y., Min, H.X., 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.
12. Wang, T., Wong, K.K., and Hong, W. (2004). A unique region of RILP distinguishes it from its related proteins in its regulation of lysosomal morphology and interaction with Rab7 and Rab34. Mol. Biol. Cell *15*, 815–826. https://doi.org/10.1091/mbc.e03-06-0413.
13. Schaub, J.R., and Stearns, T. (2013). The Rilp-like proteins Rilpl1 and Rilpl2 regulate ciliary membrane content. Mol. Biol. Cell *24*, 453–464. https://doi.org/10.1091/mbc.E12-08-0598.
14. Sobu, Y., Wawro, P.S., Dhekne, H.S., Yeshaw, W.M., and Pfeffer, S.R. (2021). Pathogenic LRRK2 regulates ciliation probability upstream of tau tubulin kinase 2 via Rab10 and RILPL1 proteins. Proc. Natl. Acad. Sci. USA *118*. e2005894118. https://doi.org/10.1073/pnas.2005894118.
15. Lara Ordónez, A.J., Fernández, B., Fdez, E., Romo-Lozano, M., Madero-Pérez, J., Lobbestael, E., Baekelandt, V., Aiastui, A., López de Munaín, A., Melrose, H.L., et al. (2019). RAB8, RAB10 and RILPL1 contribute to both LRRK2 kinase-mediated centrosomal cohesion and ciliogenesis deficits. Hum. Mol. Genet. *28*, 3552–3568. https://doi.org/10.1093/hmg/ddz201.
16. LaCroix, A.J., Stabley, D., Sahraoui, R., Adam, M.P., Mehaffey, M., Kernan, K., Myers, C.T., Fagerstrom, C., Anadiotis, G., Akkari, Y.M., et al.; University of Washington Center for Mendelian Genomics (2019). GGC Repeat Expansion and Exon 1 Methylation of XYLT1 Is a Common Pathogenic Variant in Baratela-Scott Syndrome. Am. J. Hum. Genet. *104*, 35–44. https://doi.org/10.1016/j.ajhg.2018.11.005.
17. Depienne, C., and 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.
18. Sellier, C., Buijsen, R.A.M., 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.
19. 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.e5. https://doi.org/10.1016/j.neuron.2021.03.038.
20. Sun, Q.Y., Xu, Q., Tian, Y., Hu, Z.M., Qin, L.X., Yang, J.X., Huang, W., Xue, J., Li, J.C., 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.
21. Westenberger, A., and Klein, C. (2020). Essential phenotypes of NOTCH2NLC-related repeat expansion disorder. Brain *143*, 5–8. https://doi.org/10.1093/brain/awz404.
22. 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. jmedgenet-2020-107649. https://doi.org/10.1136/jmedgenet-2020-107649.
23. Fukuda, H., Yamaguchi, D., Nyquist, K., Yabuki, Y., Miyatake, S., Uchiyama, Y., Hamanaka, K., Saida, K., Koshimizu, E., Tsuchida, N., et al. (2021). Father-to-offspring transmission of extremely long NOTCH2NLC repeat expansions with contractions: genetic and epigenetic profiling with long-read sequencing. Clin. Epigenetics *13*, 204. https://doi.org/10.1186/s13148-021-01192-5.
24. Paulson, H. (2018). Repeat expansion diseases. Handb. Clin. Neurol. *147*, 105–123. https://doi.org/10.1016/B978-0-444-63233-3.00009-9.