Jelena Pozojevic, Björn-Hergen von Holt, and Ana Westenberger*

Factors influencing reduced penetrance and variable expressivity in X-linked dystonia-parkinsonism

Abstract: X-linked dystonia-parkinsonism (XDP) is a neurodegenerative movement disorder that primarily affects adult Filipino men. It is caused by a founder retrotransposon insertion in *TAF1* that contains a hexanucleotide repeat, the number of which differs among the patients and correlates with the age at disease onset (AAO) and other clinical parameters. A recent work has identified additional genetic modifiers of age-associated penetrance in XDP, bringing to light the DNA mismatch repair genes *MSH3* and *PMS2*. Despite X-linked recessive inheritance, a minor subset of patients are female, manifesting the disease via various mechanisms such as homozygosity, imbalanced X-chromosome inactivation, or aneuploidy. Here, we summarize and discuss clinical and genetic aspects of XDP, with a focus on variable disease expressivity as a consequence of subtle genetic differences within a seemingly homogenous population of patients.

Keywords: X-linked dystonia-parkinsonism (XDP), retrotransposon insertion, repeat-length polymorphism, age-related penetrance, genetic modifiers

Introduction

X-linked dystonia-parkinsonism (XDP, DYT/PARK-*TAF1*, OMIM #314250) is a severe, neurodegenerative, and rapidly progressive movement disorder. The signs of XDP include dystonia and parkinsonism that generally start manifesting around the fourth decade of life [1, 2]. All XDP patients are of Filipino origin and descend from the common ancestor in whom the disease-causing variant initially arose. Given the position of this variant on the X-chromosome, XDP is inherited in an X-linked recessive manner and thus mostly affects men. The aim of our review is to summarize and discuss the factors contributing to the considerable variability in age-related penetrance and disease expressivity that has been observed in XDP patients, despite a seemingly identical genetic background. Furthermore, we will review the circumstances that account for the occasional presence of XDP signs in women.

Clinical characteristics of X-linked dystonia-parkinsonism

In genetically tested patients, in whom the presence of disease-causing variants has been confirmed, the median age at onset (AAO) of XDP is 40 years [3]. However, the AAO spectrum is very broad (from 20 to 67 years) and penetrance increases with age (Figure 1 Kaplan Meier Survival Analysis) [3]. The majority (>80 %) of patients present initially with focal dystonia that usually generalizes within five years after onset (*dystonic phase*) [1–3]. Nevertheless, in 14 % of patients, parkinsonism signs are the first disease manifestation [3]. With time, dystonia and parkinsonism increasingly overlap (*combined dystonia-parkinsonism phase*), and in patients surviving this disease stage the predominant clinical picture is parkinsonism (*parkinsonian phase*) [1]. XDP patients have a shortened lifespan due to dystonia-related conditions (aspiration pneumonia or starvation) or suicide, and many of them do not survive longer than 10 years after the illness onset [1].

Of note, in patients that present with dystonia at onset, craniocervical dystonia is more commonly observed (~60 %) than limb (~37 %) or truncal (~4 %) dystonia [4]. Approximately half of patients that initially exhibit craniocervical dystonia suffer from neck/shoulder dysto-

*Corresponding author: Ana Westenberger, Institute of Neurogenetics, University of Lübeck and University Hospital Schleswig-Holstein, BMF, Building 67; Ratzeburger Allee 160, 23538 Lübeck, Germany, e-mail: ana.westenberger@neuro.uni-luebeck.de
Jelena Pozojevic, Institute of Neurogenetics, University of Lübeck and University Hospital Schleswig-Holstein, BMF, Building 67; Ratzeburger Allee 160, 23538 Lübeck, Germany; and Institute of Human Genetics, University of Lübeck and University Hospital Schleswig-Holstein, Lübeck, Germany
Björn-Hergen von Holt, Institute of Medical Biometry and Statistics, University of Lübeck and University Hospital Schleswig-Holstein, Lübeck, Germany
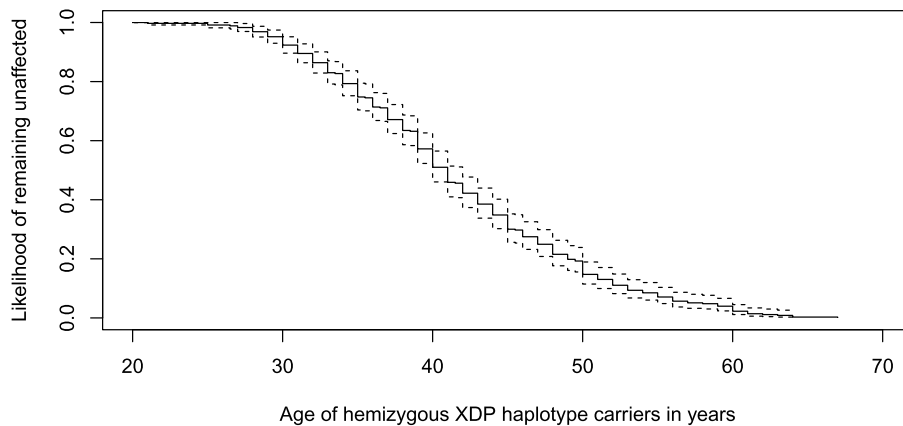
Figure 1: A Kaplan–Meier plot showing the increase in the age-related penetrance of XDP, i. e., a decrease in likelihood for men with the XDP haplotype to remain unaffected with increasing age.

nia, while 28 % and 23 % manifest blepharospasm and mouth/tongue dystonia, respectively [4]. Consistent with these findings, the initial report on 42 patients also found that first signs and symptoms occur less frequently in the legs than in other body parts (12 % head, 29 % axial musculature, 23 % upper extremities, 36 % lower extremities) [2]. Neuroanatomical changes in XDP affect the basal ganglia and cerebellum, with striking striatal atrophy and iron deposition that correlates with the degree of atrophy, disease duration, and disease severity [5–7].

Genetic basis and molecular mechanisms of XDP

XDP is endemic to the Philippines and it is particularly prevalent on Panay Island (5.74 in 100,000 individuals) [1], where a genetic founder likely originated. Indeed, all XDP patients inherited from this founder a common haplotype that consists of seven changes located within or surrounding the *TAF1* gene on the X chromosome. This XDP haplotype encompasses five disease-specific single-nucleotide changes (DSCs), a 48-bp deletion, and an SVA (SINE-VNTR-Alu) retrotransposon insertion [8, 9]. Importantly, none of these changes are situated in the protein-coding regions or present in >470 ethnically matched controls [10], leaving the identity of the actual XDP-causing variant unclear. The first experimental evidence indicating the culprit emerged from studies that edited the XDP-relevant SVA retrotransposon out of patient-derived cells. This approach led to a successful rescue of molecular phenotypes, such as the decrease in *TAF1* expression and the partial retention of

the intron containing the SVA insertion, in the edited cells [11–13].

Namely, the XDP-relevant SVA retrotransposon is inserted in intron 32 of *TAF1*, and early research showed that a neuron-specific transcript of *TAF1* was reduced in the caudate nucleus (dorsal striatum) of an XDP patient in comparison to controls without the XDP haplotype [9]. *TAF1* encodes the largest subunit of the general transcription factor TFIID that is a part of the preinitiation complex, which is indispensable for proper promoter recognition and transcription by RNA polymerase II (reviewed in [14]). Subsequent studies demonstrated a consistent downregulation of *TAF1* transcripts in various tissues and cell lines, with the *TAF1* being the only gene in the disease-linked region that shows changes in expression in XDP patients [11, 12, 15]. This effect on transcription is not surprising and has been well-documented for SVA retrotransposons in the human genome. More than 2,700 SVA elements have been found in the human genome reference sequence [16]. Although they are typically heavily methylated and thus silenced, these elements may regulate gene expression depending on the genomic context, i. e., location of their insertion [17]. Furthermore, the XDP-relevant SVA retrotransposon was shown to cause the production of an aberrant *TAF1* transcript that contains a deep intronic part of intron 32 proximal to the insertion and terminates 716 bp 5´ of the SVA [12]. The reduction in the *TAF1* expression and the retention of a part of intron 32 in some *TAF1* transcripts are currently considered relevant cellular/molecular XDP phenotypes, but their role in XDP pathogenesis is not defined and warrants further investigation.
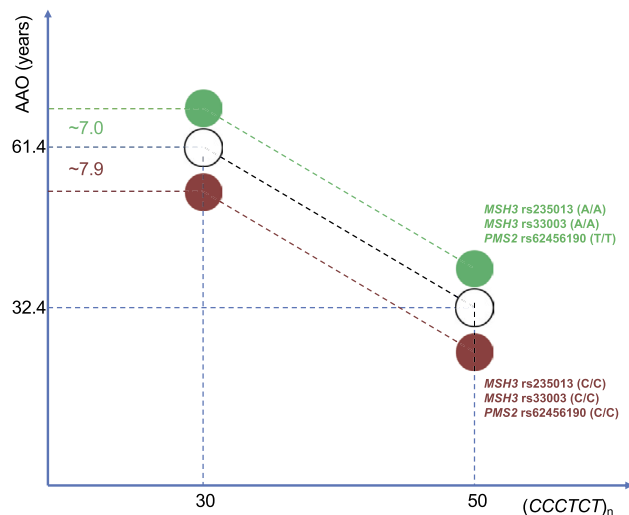
Figure 2: A diagram showing the effect of the number of $(CCCTCT)_n$ repeats, alone, on age at onset (AAO) of XDP patients (white circles) and the combined effect of the harmful alleles and the number of repeats (red circle) and protective alleles and number of repeats (green circle). The latter two decrease the AAO by 7 years and increase the AAO by 7.9 years, respectively.

Genetic modifiers of XDP and their potential modes of action

$(CCCTCT)_n$ repeat within the SVA retrotransposon

The pathogenic role of the SVA retrotransposon insertion was further substantiated by the discovery that it contains a polymorphic hexanucleotide repeat, the length of which inversely correlates with AAO in XDP patients (Figure 2) [18]. Canonical SVAs contain five consecutive regions: a polymorphic hexamer $(CCCTCT)_n$ repeat at the 5' end, followed by an Alu-like, a VNTR (variable number of tandem repeats), a SINE (short interspersed nuclear element), and a poly(A) domain at the 3' end. While the average SVA length is reported to be ~2 kb, their size can vary due to changes in copy number in their repetitive domains [19], and the XDP-specific SVA is ~2.6 kb long. In thus far investigated XDP patients, the polymorphic $(CCCTCT)_n$ repeat number ranges between 30 and 55 repeats [4]. This variability explains approximately 50 % of the variance in XDP AAO [4, 18]. Interestingly, patients presenting with parkinsonism at onset have a smaller median repeat number and later median AAO in comparison to those showing dystonia as the initial manifestation (39 vs. 41, p = 1.8×10^{-3} and 46 vs. 40, p = 8.0×10^{-5}). This relation between the repeat number and dystonia/parkinsonism at onset is indirect and determined by the AAO [4]. That is,

the number of repeats influences AAO which further affects the signs at onset, but whether a patient will initially become affected by dystonia or parkinsonism is not defined by the variability in the repeat number. In patients that present with dystonia at onset, the repeat number does not seem to be correlated with whether the disease will start in the craniocervical region, arm, leg, or trunk [4]. However, within the craniocervical region, the patients whose mouth/tongue first becomes affected have a slightly lower repeat number than those initially suffering from blepharospasm [4]. Furthermore, we described a positive correlation of the $(CCCTCT)_n$ repeat number with the severity of dystonia (i. e., the higher Burke-Fahn-Marsden Dystonia Rating Scale score) while a positive correlation with the severity of parkinsonism and cognitive dysfunction was indicated but did not reach statistical significance likely due to the low number of investigated patients [4].

The mode of action through which the hexanucleotide repeat modifies the age-related penetrance and other clinical characteristics of XDP is currently unknown. In a subset of patients, blood-derived *TAF1* mRNA levels were negatively correlated with the number of repeats, indicating that the repeat number may influence the previously observed reduction of *TAF1* expression. Additionally, and similarly to what was described for other repeat-expansion disorders (e. g., myotonic dystrophy (DM), fragile X tremor ataxia syndrome (FXTAS), C9orf72-associated amyotrophic lateral sclerosis and frontotemporal dementia (ALS/FTD)), if the XDP-relevant hexanucleotide repeat is transcribed, scenarios such as RNA foci formation [20] and repeat-associated non-ATG (RAN) translation [21] may be involved not only in modifying, but even in causing this disease. RNA foci are accumulations of RNA fragments comprised of repeated nucleotides that adopt unusual secondary structures, sequester various RNA-binding proteins, and as such become toxic to the cells. Alternatively, or in parallel, neurodegeneration may result from toxic species such as RAN-translated dipeptides.

DNA mismatch repair and repeat instability

Recently, a genome-wide association study (GWAS) was performed to uncover additional genetic factors that influence the age-related penetrance of XDP. The study found three independent genomic loci that account for an additional 25.6 % of the remaining variability, not explained by the repeats, and 13.0 % of the overall variance in AAO in XDP [22]. Two of these loci are located within or immediately adjacent to the *MSH3* gene on chromosome 5, while the third locus includes signals on chromosome 7, adjacent

to the *PMS2* gene [22]. Importantly, the combined effect of the protective alleles of the lead SNPs in each of the three regions delays XDP onset by seven years (Figure 2). The *MSH3* and *PMS2* genes code for proteins that are components of the DNA mismatch repair (MMR) cellular machinery, which is in charge of recognizing and repairing mismatched base pairs in DNA. The MMR response to larger insertion/deletion loops (up to 13 nucleotides) in particular [23] is initiated by recognition of mismatched DNA by the MutSβ complex, comprised of MSH2 and MSH3. This is followed by the recruitment of MutLα (MLH1-PMS2) heterodimers, which introduce single-strand breaks near the mismatch and enable exonucleases to degrade the mismatch-containing region. Remarkably, one of the loci in *MSH3* modifying AAO in XDP has been associated, also by GWAS, with disease progression in Huntington's disease (HD) [24], and signals in *MSH3*, *PMS2*, and other DNA maintenance genes modify AAO in HD and DM [25, 26]. The proposed mechanism of disease modification of the MMR machinery is that it recognizes longer stretches of repeats that are partially unwound and, in an effort to repair them, instigates instability, i. e., increases the length of the repeat tract. Indeed, the XDP-relevant hexanucleotide repeat displays instability in terms of intergenerational transmission and somatic mosaicism. Namely, it has been reported that in XDP families, repeat numbers tend to increase in successive generations and are more likely to be larger than in the previous generation if they are inherited from a mother than from a father [4, 18]. Furthermore, the hexanucleotide repeat exhibits somatic mosaicism in XDP patients, reflected in a higher repeat number in brain tissues compared to blood, suggesting a tissue-specific and potentially even brain-region specific effect [4, 27].

Although the exact modifiers of AAO in XDP patients, tagged by the three lead signals on chromosomes 5 and 7 are currently unknown, it is conceivable that they act by regulating expression, function, or even location of the MSH3 and PMS2 proteins. Interestingly, exon 1 of *MSH3* contains length polymorphisms, the length of which inversely correlates with *MSH3* expression and AAO in XDP, HD, and DM patients [22, 26]. The same polymorphism may have a role in regulating the levels of MSH3 in the nucleus. Specifically, nuclear localization (NLS) and nuclear export (NES) signals have been identified within MSH3 as responsible for the ability of this protein to cross the nuclear envelope [28, 29]. Furthermore, deacetylation of MSH3 by histone deacetylase 3 (HDAC3) increases repeat expansions/instability [29]. The deacetylation sites in MSH3 overlap with NLS1, and MSH3 localization is partially dependent on HDAC3 activity. Importantly, NLS1 (encoded by *MSH3* exon 2) is in close proximity to the length

polymorphism in exon 1, and it has been shown that the shortest MSH3 allele is more prone to staying in the cytoplasm in comparison to the wild type protein [28]. This is in agreement with the increase in AAO that was observed in patients carrying the shortest allele, as the absence of MSH3 from the nucleus would prevent it from introducing instability and would thus be protective. Whether this length polymorphism in exon 1 of *MSH3* is the genuine modifier behind one of the three signals seen in the XDP GWAS and what the other two are, remains to be addressed in the future studies.

Penetrance of XDP in women

In the majority of female carriers of the XDP haplotype, X chromosome inactivation (XCI) is balanced, i. e., a sufficient number of cells express solely the wildtype X chromosome for these women to remain healthy, which is a defining characteristic of the X-linked recessive inheritance. Nevertheless, occasionally female carriers of the XDP haplotype manifesting XDP are encountered in clinical practice. To date, 16 such women have been reported in the literature, and in comparison to men, they seem to have a milder phenotype, later AAO, and a milder course of the disease [30–33]. There are three possible explanations for the occurrence of XDP symptoms in females. Firstly, they may be homozygous carriers of the XDP haplotype (Figure 3A) and two such women have been hitherto described [30, 33]. Secondly, the affected women may have an imbalanced XCI, "skewed" towards the preferential expression of the affected allele. Namely, although in most women XCI ranges between 80 %:20 % and 50 %:50 % ratios of cells expressing one X chromosome vs. the other, a small percentage of females have a skewed ratio (>80 %:20 %), and thus a bias towards cells expressing one or the other X-chromosome. Interestingly, skewing increases with age, and the percentage of individuals with a >80 %:20 % XCI ratio was found to be almost three times higher among adult women (14.2 %) compared to newborn girls (4.9 %) [34]. The women whose XCI is skewed towards the preferential expression of the X chromosome that contains the SVA insertion will thus have a higher likelihood of becoming affected. Indeed, this skewing that occasionally can be extreme, accounting for less than 5 % of cells in the patient's body expressing the wildtype allele [31], is the most frequent explanation for the penetrance of XDP in women (Figure 3B). The third and the least common scenario leading to XDP expression in a female is a typical or atypical Turner syndrome,
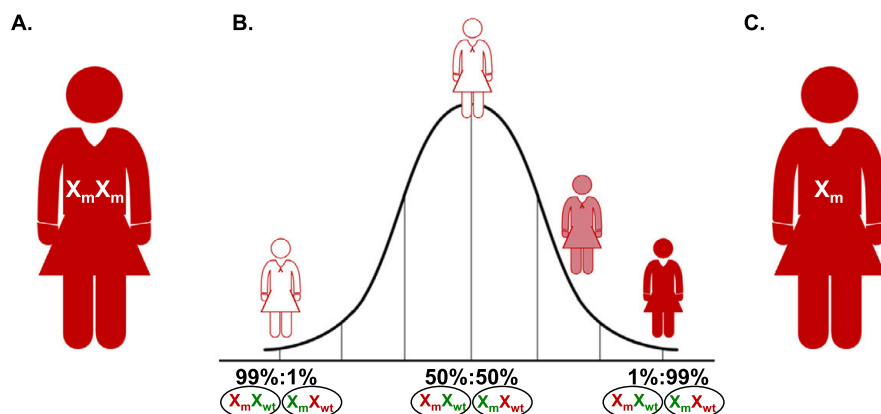
Figure 3: Explanations for penetrance of XDP in women. A. A homozygous carrier of the XDP haplotype. B. A Bell curve showing the effect of X-chromosome inactivation (XCI) on the penetrance of XDP in heterozygous female XDP-haplotype carriers. X_m – affected X chromosome (i. e., X chromosome containing the XDPhaplotype), X_{wt} wildtype X chromosomes. The silenced X chromosome is shown in red and the one that is expressed in green. The percentages below the curve are the ratios of cells with silenced affected vs. unaffected X chromosome. White female figures – unaffected heterozygous carriers; light and dark red female figures, mildly or severely affected female XDP patients, respectively. C. A woman with X chromosome monosomy (45, X), harboring only the affected X chromosome in all her cells.

i. e., complete (45,X) or mosaic (45,X/46,XX) X chromosome monosomy, resulting in a loss of the wildtype X chromosome in all or some cells of the affected woman (Figure 3C). Under these circumstances, the XDP-relevant genetic setting is comparable between a typical male XDP patient and a Turner syndrome XDP patient. Of note, atypical Turner syndrome, in particular, may go undiagnosed because of the mild, unremarkable phenotype is appreciated only upon XDP-relevant genetic testing [32]. Interestingly, in terms of X-chromosome aneuploidies, Klinefelter syndrome, a relatively frequent (1:500 to 1:1,000 males) condition characterized by a presence of an additional X chromosome in a male genotype (47,XXY), may be protective and reduce XDP penetrance in individuals heterozygous for the XDP haplotype with balanced XCI.

Concluding remarks

Even though all XDP patients can track their ancestry to the Panay island in the Philippines and share the common haplotype, there are subtle differences among this seemingly genetically homogenous population. In addition to the polymorphic hexanucleotide repeat within the SVA, modifiers in MMR genes determine age-related disease penetrance. Combined, these genetic factors explain approximately 63 % of AAO variability in XDP patients, indicating that additional modifiers remain to be identified. Understanding the molecular mechanisms of disease modification has a multifaceted benefit for XDP patients that includes improved genetic counseling, more

efficient selection and enrollment of patients in clinical studies and trials, and may even lead to the development of specific therapies. While this important newly acquired knowledge is applicable to all the mutation carriers, additional factors that might modify disease expressivity in isolated cases should be considered. This is illustrated by the rare female patients reported to manifest XDP due to epigenetic and copy-number variation, including imbalanced XCI and X-chromosome monosomy.

References

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

[2]   Lee LV, Kupke KG, Caballar-Gonzaga F et al. The phenotype of the X-linked dystonia-parkinsonism syndrome: An assessment of 42 cases in the Philippines. Medicine. 70:179–87.

[3]   Pauly MG, Ruiz López M, Westenberger A et al. Expanding Data Collection for the MDSGene Database: X-linked dystonia-parkinsonism as use case example. Mov Disord. 2020;35:1933–8.

[4]   Westenberger A, Reyes CJ, Saranza G et al. A hexanucleotide repeat modifies expressivity of X-linked dystonia parkinsonism. Ann Neurol. 2019;85(6):812–22.

[5] Pasco PMD, Ison CV, Muñoz EL et al. Understanding XDP through imaging, pathology, and genetics. Int J Neurosci. 2010;121:12–7.

[6] Hanssen H, Heldmann M, Prasuhn J et al. Basal ganglia and cerebellar pathology in X-linked dystonia-parkinsonism. Brain. 2018;141(10):2995–3008.

[7] Hanssen H, Prasuhn J, Heldmann M et al. Imaging gradual neurodegeneration in a basal ganglia model disease. Ann Neurol. 2019;86:517–26.

[8] Nolte D, Niemann S, Müller U. Specific sequence changes in multiple transcript system DYT3 are associated with X-linked dystonia parkinsonism. Proc Natl Acad Sci USA. 2003;100(18):10347–52.

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

[10] Domingo A, Westenberger A, Lee LV et al. New insights into the genetics of X-linked dystonia-parkinsonism (XDP, DYT3). Eur J Hum Genet. 2015;23:1334–40.

[11] Domingo A, Amar D, Grütz K et al. Evidence of TAF1 dysfunction in peripheral models of X-linked dystonia-parkinsonism. Cell Mol Life Sci. 2016;73(16):3205–15.

[12] Aneichyk T, Hendriks WT, Yadav R et al. Dissecting the causal mechanism of X-linked dystonia-parkinsonism by integrating genome and transcriptome assembly. Cell. 2018;172:897–909.e21.

[13] Rakovic A, Domingo A, Grütz K et al. Genome editing in induced pluripotent stem cells rescues TAF1 levels in X-linked dystonia-parkinsonism. Mov Disord. 2018;33:1108–18.

[14] Bhuiyan T, Timmers HTM. Promoter Recognition: Putting TFIID on the Spot. Trends Cell Biol. 2019;29(9):752–63.

[15] Ito N, Hendriks WT, Dhakal J et al. Decreased N-TAF1 expression in X-linked dystonia-parkinsonism patient-specific neural stem cells. Dis Model Mech. 2016;9(4):451–62.

[16] Wang H, Xing J, Grover D et al. SVA elements: A hominid-specific retroposon family. J Mol Biol. 2005;354(4):994–1007.

[17] Elbarbary RA, Lucas BA, Maquat LE. Retrotransposons as regulators of gene expression. Science. 2016;351(6274):aac7247.

[18] Bragg DC, Mangkalaphiban K, Vaine CA et al. Disease onset in X-linked dystonia-parkinsonism correlates with expansion of a hexameric repeat within an SVA retrotransposon in TAF1. Proc Natl Acad Sci USA. 2017;114:E11020-8.

[19] Gianfrancesco O, Geary B, Savage AL et al. The role of SINE-VNTR-Alu (SVA) retrotransposons in shaping the human genome. Int J Mol Sci. 2019;20(23):5977.

[20] Zhang N, Ashizawa T. RNA toxicity and foci formation in microsatellite expansion diseases. Curr Opin Genet Dev. 2017;44:17–29.

[21] Cleary JD, Ranum LPW. Repeat-associated non-ATG (RAN) translation in neurological disease. Hum Mol Genet. 2013;22(R1):R45–51.

[22] Laabs BH, Klein C, Pozojevic J et al. Identifying genetic modifiers of age-associated penetrance in X-linked dystonia-parkinsonism. Nat Commun. 2021;12(1):3216.

[23] Palombo F, Iaccarino I, Nakajima E et al. hMutSβ, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. Curr Biol. 1996;6:1181–4.

[24] Moss DJH, Pardiñas AF, Langbehn D et al. Identification of genetic variants associated with Huntington's disease progression: a genome-wide association study. Lancet Neurol. 2017;16(9):701–11.

[25] Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium. CAG repeat not polyglutamine length determines timing of Huntington's disease onset. Cell. 2019;178(4):887–900.e14.

[26] Flower M, Lomeikaite V, Ciosi M et al. MSH3 modifies somatic instability and disease severity in Huntington's and myotonic dystrophy type 1. Brain. 2019;142(7):1876–86.

[27] Reyes CJ, Laabs B-H, Schaake S et al. Brain regional differences in hexanucleotide repeat length in X-linked dystonia-parkinsonism using nanopore sequencing. Neurol Genet. 2021;7(4):e608.

[28] Tseng-Rogenski SS, Munakata K, Choi DY et al. The human DNA mismatch repair protein MSH3 contains nuclear localization and export signals that enable nuclear-cytosolic shuttling in response to inflammation. Mol Cell Biol. 2020;40(13):e00029-20.

[29] Williams GM, Paschalis V, Ortega J et al. HDAC3 deacetylates the DNA mismatch repair factor MutSβ to stimulate triplet repeat expansions. Proc Natl Acad Sci USA. 2020;117(38):23597–605.

[30] Evidente VGH, Nolte D, Niemann S et al. Phenotypic and molecular analyses of X-linked dystonia-parkinsonism ("lubag") in women. Arch Neurol. 2004;61(12):1956–9.

[31] Domingo A, Lee LV, Brüggemann N et al. Woman with X-linked recessive dystonia-parkinsonism: Clue to the epidemiology of parkinsonism in Filipino women? JAMA Neurol. 2014;71(9):1177–80.

[32] Westenberger A, Rosales RL, Heinitz S et al. X-linked dystonia-parkinsonism manifesting in a female patient due to atypical turner syndrome. Mov Disord. 2013;28(5):675–8.

[33] Weissbach A, Bäumer T, Rosales R et al. Neurophysiological fingerprints of X-linked dystonia-parkinsonism: A model basal ganglia disease. Mov Disord. 2015;30(6):873–5.

[34] Amos-Landgraf JM, Cottle A, Plenge RM et al. X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. Am J Hum Genet. 2006;79(3):493–9.