Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

Myotonic dystrophy: disease repeat range, penetrance, age of onset, and relationship between repeat size and phenotypes

Kevin Yum[1], Eric T Wang[3] and Auinash Kalsotra[1,2]

Myotonic dystrophy (DM) is an autosomal dominant neuromuscular disease primarily characterized by myotonia and progressive muscle weakness. The pathogenesis of DM involves microsatellite expansions in noncoding regions of transcripts that result in toxic RNA gain-of-function. Each successive generation of DM families carries larger repeat expansions, leading to an earlier age of onset with increasing disease severity. At present, diagnosis of DM is challenging and requires special genetic testing to account for somatic mosaicism and meiotic instability. While progress in genetic testing has been made, more rapid, accurate, and cost-effective approaches for measuring repeat lengths are needed to establish clear correlations between repeat size and disease phenotypes.

Addresses
[1] Department of Biochemistry, University of Illinois, Urbana-Champaign, USA
[2] Institute of Genomic Biology, University of Illinois, Urbana-Champaign, USA
[3] Department of Molecular Genetics & Microbiology, Center for Neurogenetics, University of Florida, Gainesville, FL 32610, USA

Corresponding authors: Wang, Eric T (eric.t.wang@ufl.edu), Kalsotra, Auinash (kalsotra@illinois.edu)

Current Opinion in Genetics & Development 2017, **44**:30–37

This review comes from a themed issue on **Molecular and genetic bases of disease**

Edited by **Nancy Bonini**, **Edward Lee** and **Wilma Wasco**

http://dx.doi.org/10.1016/j.gde.2017.01.007

Introduction

DNA repeat expansions are responsible for more than 20 inherited neurological disorders—some of these include Huntington's disease, fragile X syndrome, spinal and bulbar muscular atrophy, as well as the most common form of familial amyotrophic lateral sclerosis [1]. In multiple repeat diseases, repeat length is correlated to disease severity and age of onset [2], yet molecular pathways that go awry due to expanded repeats can differ. Studies of myotonic dystrophy (dystrophia myotonica, DM) first demonstrated the concept that microsatellite repeats in noncoding regions can be transcribed into pathogenic RNAs [3]. Expansions can occur in the germline, leading to genetic anticipation across multiple generations, and can also occur somatically during various stages of human development with preferences for distinct tissues, ages, genders, and populations [4–7]. Furthermore, the rate of expansion in somatic cells can vary within the same tissue [8].

Overview of myotonic dystrophy

Myotonic dystrophy exists in two clinically and molecularly defined forms: myotonic dystrophy type 1 (DM1), also known as Steinert's disease; and myotonic dystrophy type 2 (DM2), also known as proximal myotonic myopathy, both of which are inherited in an autosomal dominant fashion [9]. DM1 is caused by a CTG expansion in the 3′ untranslated region of the dystrophia myotonica protein kinase (DMPK) gene on chromosome 19q13 [10,11], while DM2 is caused by a CCTG expansion located within intron 1 of the cellular nucleic-acid-binding protein (CNBP, formerly ZNF9) gene on chromosome 3q21 [12].

A healthy individual with normal DMPK alleles has 5–37 repeats (35 has also commonly been used as an upper threshold for normal repeat length [13]) [14]. DM1 patients who have repeats between 38 and 50 are said to have a "pre-mutation" allele and can be asymptomatic throughout their lifetime. However, they are at increased risk of having children with larger repeats [15]. Penetrance tends to grow as repeat length increases, but extreme variability in penetrance of specific symptoms exists in the patient population [14]. Somatic mosaicism and intergenerational instability are biased towards expansion in DM1 [4], although contraction can rarely occur. It is estimated that a decrease in the CTG repeat size during transmission from parents to child is about 6.4%, most frequently during paternal transmissions [16]. Children of DM1 parents typically inherit repeat lengths considerably larger than those present in the transmitting parent, the phenomenon known as "anticipation," where disease severity increases and age of onset decreases in successive generations. Up to 5% of DM1 patients have interrupted repeats, in which the CTG repeat tract contains GGC, CCG, or CTC repeats [17,18]. Some of these interruptions have been associated with stabilization of the CTG repeat tract length [19].

The repeat expansion of DM2 in intron 1 of CNBP is found within the context of a complex $(TG)_n(TCTG)_n(CCTG)_n$ sequence. While non-pathogenic alleles contain up to 26 repeats, the range of repeats in patients is extremely broad, with measurements from 75 to 11 000 units (on average 5000) [12]. Unlike DM1, the size of the repeat DNA expansion in DM2 does not correlate with age of onset or disease severity [20]. Furthermore, individuals homozygous for repeat expansions have clinical features indistinguishable from that of their heterozygous siblings [21]. Phenotypes and anticipation in DM2 are almost always milder than DM1, and DM2 lacks the congenital form [22].

The combined prevalence of DM1 and DM2 is approximately 1 in 8000 (12.5 per 100 000), but this is likely an underestimate because of difficulty in clinical identification of minimally affected individuals [7]. Although DM2 is generally rarer than DM1, recent epidemiological data in Germany and Finland suggest that DM2 occurs more frequently than previously observed [23]. Similarly, the prevalence of DM1 can vary widely: in Taiwan, approximately 0.5 in 100 000 people are affected; while in the United Kingdom, the number can range from 7.1 to 10.6 in 100 000 [24]. Different factors could play a role in such variations: for instance, a founder effect is assumed to have increased the prevalence of DM1 to 1 in 500 in the Saguenay–Lac-Saint-Jean region of Northeastern Quebec [25].

Despite these key differences, DM1 and DM2 share several hallmark clinical features such as myotonia, cataracts, and cardiac conduction defects [26]. The fact that two independent mutations cause similar disease pathology has led to the RNA toxicity hypothesis where the expanded repeat-containing RNAs form ribonuclear foci that sequester and disrupt the normal activities of RNA binding proteins belonging to the MBNL and CELF families [27] (for more details, see Ashizawa's review on "RNA foci" and Thornton's review on "DM: approaches to therapy" in the same issue, as well as Ref. [28]). In this review, we discuss mechanisms of repeat expansion, approaches for measuring repeat lengths, and the relationships between repeat length and phenotypes in DM.

Mechanisms of repeat expansion in mitotic and post-mitotic tissues

Several molecular mechanisms for repeat instability have been proposed, mainly in the context of DNA replication, recombination, transcription and/or repair (Box 1). Most of these mechanisms involve folding of microsatellite repeats into an unusual secondary structure, kinetically trapping the otherwise unstable DNA repeats [29]. In the case of DM, $(CTG)_n(CAG)_n$ and $(CCTG)_n(CAGG)_n$ repeats form hairpin-like secondary structures, which are stabilized by both Watson–Crick (WC) and non-WC base pairs [30,31].

Proof of principle studies from yeast have demonstrated that repeat instability can be based on replication fork stalling and restart [32,33], ruling out the classic model of strand slippage for DM [34]. In this model, the formation of a stable secondary structure during lagging strand synthesis could stall a DNA polymerase, slowing down the overall replication fork progression as the lagging and leading strand syntheses are coordinated. To minimize the stalling of replication fork, DNA polymerase can skip an Okazaki fragment to resume lagging strand synthesis (contraction pathway) or promote fork reversal (for more details on the fork reversal mechanism, please see the review by Neelsen and Lopes [35•]) to generate a structure-prone single-stranded repeat extension at the 3′ end of the leading strand. If the 3′ repetitive hairpin persists when the replication restarts, repeats can expand. Recombination can also account for repeat instability in mitotically dividing cells. In bacteria, longer repeats increase the rate of recombination [36,37] while in yeast, CTG repeats cause chromosomal breakage [38]. As Mirkin described in his review [39], one possible mechanism for recombination-based instability is that repeats promote the double-strand breaks in DNA causing the invasion of fragments into sister chromatids.

The transcription and repair models of repeat instability can likewise account for expansion in both mitotic and post-mitotic cells. During transcription, the formation of slip-outs on either strand can stall RNA polymerase II, facilitating the transcription-coupled repair. Depending on the location of the excision, subsequent patch repair could lead to expansion or contraction [40]. Similarly, studies on transgenic mouse models of Huntington's disease and DM have demonstrated that loss of MSH2/MSH3 mismatch repair proteins can decrease the frequency of repeat expansion [41–43]. This discovery has led to a theory that the MSH2/MSH3 complex can stabilize the secondary structure and prevent the flip removal by FEN1, leading to expansion during DNA repair. Although this theory is highly supported in a yeast model [44], it is less clear in a mouse model where repeat instability was unaltered in Fen1-knockout mice [45].

As disease symptoms in DM are most prominent in post-mitotic tissues such as the heart, skeletal muscle, and central nervous system (CNS), it is thought that DNA repair-dependent mechanisms, and potentially transcription-coupled nucleotide excision repair, may drive repeat instability in these tissues. For more details on tissue-specific DNA repair mechanisms of repeat instability, see the review by Dion [46•].

Diagnosis and laboratory methods to measure repeat lengths

A wide variety of DM symptoms can bring patients to the clinic, including myotonia, muscle weakness, cardiac arrhythmias, hypersomnia, gastrointestinal (GI) tract