

Myotonic dystrophy: the correlation of (CTG) repeat length in leucocytes with age at onset is significant only for patients with small expansions

Marion G Hamshere, Helen Harley, Peter Harper, J David Brook, John F Y Brookfield

Abstract

Myotonic dystrophy (DM) was the first of a group of diseases to be identified for which the genetic basis is the expansion of a triplet repeat. Myotonic dystrophy also exhibits anticipation, in which the disease worsens through successive generations. These two features have led many groups to analyse whether a significant negative correlation between triplet repeat length and severity of disease exists. However, the recent molecular finding that two distinct subsets of classically affected DM patients exist, those who export expansion derived DMPK RNA and those who do not, led us to question whether combining data from these two sets of patients is statistically valid. We found that although patients with small expansions showed a significant correlation between age at onset and triplet repeat length, those with larger expansions did not. The theoretical triplet repeat size, which separated the two groups, was also deduced.

(*J Med Genet* 1999;36:59-61)

Keywords: myotonic dystrophy; triplet repeats; age at onset

Department of
Genetics, Queen's
Medical Centre,
University of
Nottingham,
Nottingham NG7 2UH,
UK

M G Hamshere
J D Brook
J F Y Brookfield

Centre for Medical
Genetics, City Hospital
NHS Trust,
Nottingham NG6 1PB,
UK

M G Hamshere
J D Brook

Institute of Medical
Genetics, University of
Wales College of
Medicine, Cardiff CF4
4XN, UK
H Harley*
P Harper

Correspondence to:
Dr Hamshere.

*Present address: Welsh
School of Pharmacy,
University of Wales Cardiff,
Cardiff CF1 3XF, UK.

Received 11 July 1997
Revised version accepted for
publication 18 June 1998

Myotonic dystrophy (DM) is associated with a triplet expansion in the 3' untranslated region of a protein kinase gene DMPK.¹ It has become widely accepted that the length of the triplet repeat in leucocytes correlates with the age at onset of symptoms in myotonic dystrophy, with larger expansions leading to an earlier age at onset than smaller expansions. Recent experiments have led us to question this.² We found that for fibroblast patient cell lines with triplet repeat sizes of 1.2 kb and above, transcripts from the DM allele of DMPK were absent from the cytoplasm, whereas in a cell line with an expansion of only 0.24 kb the DM transcripts were processed normally. This indicated that a critical repeat size existed in fibroblasts, above which RNA was not exported to the cytoplasm. If this effect is mirrored in other tissues, then a threshold size might be expected, above which the severity of disease (reflected by age at onset) shows no significant correlation with the length of the triplet repeat.

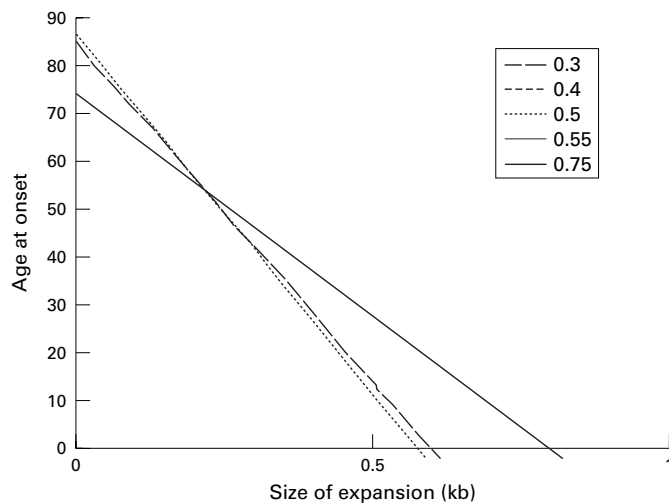
In view of these molecular findings we have reanalysed the relationship between reported age at onset in DM patients and triplet size in leucocytes by dividing these data into three subsets: those with repeat sizes below 0.24 kb (below the threshold), between 0.24 and 1.2 kb (a mixed group, some below and some above

the threshold), and above 1.2 kb (above the threshold). This allows us to determine whether the significant correlation of the data set reported previously as a whole³ is in fact the result of a significant correlation in only a subset of patients. Patients in all groups exhibited a range of symptoms including cataracts, muscle weakness, and myotonia.⁴ It is valid to differentiate, and exclude, the data from congenitally affected patients as they exhibit different and additional (mental retardation) symptoms to the classical myotonia of adult onset DM.⁴ However, the clinical findings for minimally affected patients overlap those of more severely affected classical DM cases.

Analysis of 62 patients with expansions greater than 1.2 kb shows that for this group there was no significant correlation between the age at onset and the size of their triplet repeat. The mean age at onset for these patients is 21.7 years. Non-parametric analysis of data from patients with expansion sizes at the two extremes of this group also indicates that there is no significant difference between their age at onset (by Mann-Whitney U). Therefore, both methods indicate that the age at onset for patients with expansion sizes greater than 1.2 kb is not determined by the length of their triplet repeat in leucocytes. This analysis indicates that the original data set as a whole consists of at least two groups; one of these groups (patients with repeat sizes of 1.2 kb and greater) shows no significant correlation between triplet repeat length and age at onset of DM. The subset of patients with expansion sizes less than 0.24 kb contains too few data points for statistical analysis, and analysis of the data from patients with expansion sizes of less than 1.2 kb as a whole is not valid, as these are predicted to contain heterogeneous data (that is, some patients with expansions below and some above the threshold).

In order to establish the triplet repeat size at which there is a transition (a theoretical threshold) from one group (which shows a significant correlation) to another (which does not), serial regression analyses were undertaken, with age at onset regressed on size of repeat. The initial regression, in this analysis, is based on data from 23 patients who have expansions of 0.3 kb or less. The incremental addition of data from other patients, whose expansions are in the next size category indicates, by a change in the slope of the line, the point at which data points from a different group are now also included (fig 1).

It is clear that when data from patients with expansion sizes between 0.3 and 0.4 kb are



No of data points used	Size of expansion (kb)	Intercept (age at onset)	Slope age/kb
23	<or = 0.3	85	-142
25	<or = 0.4	86	-150
35	<or = 0.5	75	-92.9
36	<or = 0.55	74	-92.4
37	<or = 0.75	74	-92.2

Figure 1 Graph showing the age at onset regressed on size of triplet expansion in patients with adult onset myotonic dystrophy. The regression lines for 0.5, 0.55, and 0.75 are coincident. All sizes were obtained by Southern blot analysis, although PCR was also used to confirm the length of smaller expansions. Total numbers of patients in each group, data groupings, and equations from regression analysis are shown in the table below.

added to those from patients with expansions of less than 0.3 kb there is no change in the slope of the line. However, when data from patients with expansions between 0.4 and 0.5 kb are included in the data set, the slope of the line is shifted. Therefore, the theoretical threshold appears to be within the group with expansions between 0.4 and 0.5 kb and the transition point can be calculated. Below 0.4 kb the equation for regression analysis is $y = -150x + 86$. Above 1.2 kb (where no significant correlation between age at onset and triplet repeat length is seen), the equation is $y = 21.7$. The theoretical transition point is the intersect of these two lines, a value of 0.43 kb.

Patients with expansions between 0.43 and 1.2 kb could be either from the same group as those with expansions greater than 1.2 kb or could represent a transition group. By non-parametric analysis, comparison of data from patients with expansions between 0.43 and 1.2 kb to data from patients with expansions between 3.0 and 4.75 kb show that there is no significant difference in the median ages for the two groups (median age at onset of 25 and 23.5, respectively), indicating that they probably formed part of the same group. In contrast, there is a significant difference between the ages of the 0.43 to 1.2 kb group (median age at onset of 25) and those from the group with expansions of less than 0.43 kb (median age at onset of 55) ($p < 0.001$ by Mann-Whitney).

These findings indicate that when the CTG triplet expansion in leucocytes is greater than 1.2 kb, the absolute size of the repeat will not

be a good indicator of the expected age at onset of myotonic dystrophy symptoms. Indeed, this value may be as low as 0.43 kb. An alternative explanation may be that these findings in leucocytes do not reflect the situation in other tissues, such as muscle, for which a significant correlation could exist, because of somatic heterogeneity. However, as a correlation has been noted for the size of expansion in leucocytes and muscle⁵ and between leucocytes and fibroblasts (JDB unpublished data), this explanation is less likely.

The identification of a second gene (DMAHP) immediately downstream of the expansion mutation, for which the (CTG) expansion may disrupt a CpG island, led to the possibility that expression of this gene may also be disrupted in DM.⁶ Recently, expression from the expansion allele of DMAHP has been shown to be reduced in cell lines and muscle biopsies from DM patients.^{7,8} However, a more comprehensive study into the relationship between triplet repeat length and the level of expression from the expansion derived DMAHP gene will be required before the role of DMAHP can be related to this lack of significant correlation between age at onset and triplet repeat length for patients with triplet repeat sizes above 0.43 kb.

It is unclear how the additional symptoms found in congenital DM could arise if there is no correlation between the length of the triplet repeat when above 0.43 kb and the severity of symptoms in classically affected DM patients. Recently, however, a hypermethylated site, close to the triplet repeat expansion, has been found to be associated with patients with congenital but not classical DM.⁹ This has been proposed as a second genetic factor causally related to the congenital form of the disease.⁹ In this model, hypermethylation of a site close to the triplet repeat could lead to an alteration in the expression of another gene (for example, DMAHP) for which loss of function is only detrimental during early development.

In the light of the findings reported here, it will be interesting to reanalyse other data relating age at onset and triplet repeat length, and to assess how an alteration in the level of DMAHP may relate to the size of the triplet repeat.

This work was funded by the Muscular Dystrophy Association (USA) and the University of Nottingham.

- 1 Brook JD, McCurrach ME, Harley HG, et al. Molecular-basis of myotonic-dystrophy - expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein-kinase family member. *Cell* 1992;68:799-808.
- 2 Hamshere MG, Newman EE, Alwazzan M, Athwal BS, Brook JD. Transcriptional abnormality in myotonic dystrophy affects DMPK but not neighbouring genes. *Proc Natl Acad Sci USA* 1997;94:7394-9.
- 3 Harley HG, Rundle SA, MacMillan JC, et al. Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission in myotonic dystrophy. *Am J Hum Genet* 1993;52:1164-74.
- 4 Harper P. *Myotonic dystrophy*. London: Saunders, 1989.
- 5 Anvret M, Ahlberg G, Grandell U, et al. Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy. *Hum Mol Genet* 1993;2:1397-400.
- 6 Boucher CA, King SK, Carey N, et al. A novel homeodomain-encoding gene is associated with a large CpG island interrupted by the myotonic-dystrophy unstable (CTG)(n) repeat. *Hum Mol Genet* 1995;4:1919-25.
- 7 Thornton CA, Wymer JP, Simmons Z, McClain C, Moxley RT. Expansion of the myotonic dystrophy CTG repeat

- reduces expression of the flanking *DMAHP* gene. *Nat Genet* 1997;16:407-9.
- 8 Klesert TR, Otten AD, Bird TD, Tapscott SJ. Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of *DMAHP*. *Nat Genet* 1997;16:402-6.
- 9 Steinbach P, Glaser D, Vogel W, Wolf M, Schwemmle S. The *DMPK* gene of severely affected myotonic dystrophy patients is hypermethylated proximal to the largely expanded CTG repeat. *Am J Hum Genet* 1998;62:278-85.