

Editorial

Instability versus predictability: the molecular diagnosis of myotonic dystrophy

In this issue of the Journal a series of papers assesses the utility of molecular analysis of the mutation within the presumed myotonic dystrophy (DM) gene in the management of this disorder. A detailed correlation of phenotype and genotype will require larger and more comprehensive studies, but already it is clear that molecular analysis will be the basis for the presymptomatic and prenatal diagnosis of DM in the future. This editorial reviews the development of these molecular studies and their application in clinical practice.

Clinical genetics of DM

DM is the most common form of muscular dystrophy among Caucasians, having a disease prevalence four times that of Duchenne muscular dystrophy (distributed over both sexes) and an incidence of 1 in 8000.¹ A clinician's perspective of DM depends on his/her speciality. For a neurologist, the classical presentation of a distinct pattern of muscle weakness in association with myotonia in early adulthood represents a straightforward diagnosis. Paediatricians think of DM in terms of the neonate with hypotonia, facial diplegia, and respiratory difficulties, the features of congenital DM. Astute ophthalmologists may recognise systemic features in patients presenting for cataract surgery. For the clinical geneticist, DM is the archetypal dominant disorder with wide variation of expression and multisystem involvement. Two of its clinical features, increasing severity in successive generations (anticipation) and exclusive maternal inheritance in cases of congenital DM, have been difficult to explain in terms of classical mendelian genetics. Indeed, whether anticipation is a biological reality or simply ascertainment bias has been a matter of debate for over 80 years.²

Before the advent of DNA markers, assessment of at risk subjects depended on clinical, EMG, and slit lamp examinations. It was widely accepted that the distinctive EMG changes and polychromatic lens crystals were highly specific for DM, although it was not known at what age carrier status could be excluded on the basis of a negative examination.¹ With the application of linked DNA markers, it became clear that 8 to 10% of carriers of reproductive age were normal when evaluated clinically and by EMG.³ The main

dilemma in counselling was predicting the severity of the disease for an individual person. In particular, it was difficult to predict which women were at risk of having congenitally affected babies. For any female carrying the mutation, the overall risk of having a congenitally affected child is approximately 10%, with a further 10% dying in the perinatal period, and 30% developing DM at a later stage.¹ However, if the mother has clear evidence of DM at the time of the pregnancy, the risk of having a congenitally affected baby is approximately 40%; this risk probably rises to 50% once an affected baby has been born.⁴ Conversely, a female carrier with no evidence of the disease is at comparatively low risk of having a baby affected at birth.

Molecular genetics of DM

Linkage studies mapped the DM gene to chromosome 19, and this genetic localisation ultimately led to the recent identification of an unstable (CTG)_n repeat as the molecular basis of the mutation.⁵⁻⁷ The repeat is located in the 3' non-coding region of the myotonin protein kinase gene.⁸⁻¹⁰ The wild type allele has been reported to have five to 30 copies of this repeat, but affected subjects have between 50 and 2000 copies. Papers in this issue of the Journal increase the upper limit of the normal range to 37 copies.^{11,12} An increase in (CTG)_n copy number (or amplification of the repeat) is presumed to modify the expression of the gene. On normal chromosomes the copy number is stable during meiosis and mitosis, and the (CTG)_n repeat segregates simply as a polymorphic locus. However, in families with DM, the copy number increases during successive generations. Initial studies have suggested that the degree of amplification correlates with the severity of the phenotype, thereby providing a molecular correlate of anticipation.^{13,14} Furthermore, somatic mosaicism may be apparent with different degrees of amplification in different cells within a subject.^{6,9,10,13}

Thus, the (CTG)_n repeat shows the features of hereditary unstable DNA initially characterised in the fragile X mutation.¹⁵ In both conditions the copy number on the mutant allele tends to increase in successive generations; occasional small decreases in copy number have been noted in the fragile X

syndrome,¹⁶⁻¹⁸ and a similar phenomenon may occur in DM. In both conditions an increase in copy number can occur in the absence of recombination. The sequence of the DM repeat (CTG)_n differs from that found at the fragile X (CCG)_n, and some details of the mutation dynamics^{10,17,18} may well be different. The only other human mutation identified to date as being an amplified triplet repeat is that found at the androgen receptor gene in spinobulbar muscular atrophy (SBMA, Kennedy syndrome).¹⁹ The SBMA repeat sequence is the same as that found in DM, but the distribution of (CTG)_n copy numbers in the normal population differs at the two loci.²⁰ In contrast to the fragile X and DM repeats, the SBMA repeat occurs within the coding region of a gene, and there is no evidence of the instability noted in the other two conditions (but see note added in proof). As an aside, it would be to the advantage of the whole genetic community if a simple method for describing the degree of amplification of these various repeats was generally accepted.

Correlation of genotype and phenotype

If mutation analysis is to be used in the management of families with DM, it is essential that the relationship between different mutations and the phenotype be clearly defined. Two elements of the genotype/phenotype correlation must be considered separately. On one hand, studies of a large number of affected subjects may provide unequivocal evidence of the general correlation between (CTG)_n copy number and phenotype, and hence lend support to a particular hypothesis regarding molecular pathogenesis. The situation in genetic counselling is different, for here one is concerned to give accurate and precise advice to one person. The instability of the (CTG)_n repeat both in an individual person and within a family necessarily places a restriction on the precision of the advice that can be given. Papers in this issue of the Journal address both aspects of mutation analysis.

The association of increased (CTG)_n copy number and DM has been confirmed in numerous studies of primarily Caucasian populations.^{5-10,13,14} The same association has been noted in Japanese patients²¹ and, in this issue of the Journal, Melchionda *et al*²² extend the observations to Italian and Spanish populations. The amplification is detectable by Southern blot in over 70% of cases, but in a significant proportion the amplification can only be detected by PCR.^{9,20,22} To date, in excess of 700 affected subjects have been assessed by both Southern blot and PCR,^{13,14,22} and an increase in (GTG)_n copy number has been documented in over 99%. Over 1300 normal chromosomes have been studied, and all have copy numbers of less than 40.^{8-10,13} Mahadevan *et al*⁹ reported two pedigrees where affected subjects did not have evidence of amplification. In this issue of the Journal, Brunner *et al*¹¹ report a third DM pedigree where the affected subjects do not show amplification of the (CTG)_n repeat. Detailed clinical

and molecular evaluation of these families is awaited with great interest.

Predictive testing using mutation analysis

It is clear that analysis of the (CTG)_n repeat is the investigation of choice in certain diagnostic settings. As highlighted by Reardon *et al*²³ and Brunner *et al*¹¹ in this issue, molecular studies can be invaluable when evaluating an isolated proband with equivocal signs of DM, or when evaluating the apparently normal parents of a classically affected person. It must be remembered that the relatively small increases in (CTG)_n copy number found in these instances are best detected by PCR. Presymptomatic diagnosis of carriers by mutation analysis is quicker and less liable to error than prediction on the basis of flanking polymorphic loci, clinical examination, EMG, or slit lamp examination.

The degree of amplification correlates with the severity of the disease as measured by age of onset.^{13,14} In this issue of the Journal, Hunter *et al*²⁴ report the correlation of age of onset of symptoms with (CTG)_n copy number in a further large series of patients and affected parent/child pairs. There was an inverse relationship between the age of onset of DM and (CTG)_n copy number, and an increase in copy number in successive generations was associated with earlier age of onset in offspring. Difficulties arise when applying these data in genetic counselling as the relationship is not absolute. In general, it is not possible to predict the age of onset of the disease for a subject on the basis of the (CTG)_n copy number. For any given degree of amplification there is a wide range of disease severity. The only clear distinction that can be made is between congenital DM (with (CTG)_n copy numbers in excess of 500; amplification readily detected on Southern blot) and the minimal form of the disease (with copy numbers between 50 and 100; amplification not detected on Southern blot).^{13,24}

Estimating the risk of congenital DM

The clinical study of Koch *et al*⁴ clearly showed that evidence of DM in more than one body system identifies women at very high risk of having congenitally affected babies. If a mother has multisystem DM at the time of her pregnancy and her baby inherits the DM mutation, the risk of the baby being congenitally affected is approximately 80%.⁴ Apart from showing that a baby has indeed inherited the DM mutation, can analysis of the (CTG)_n repeat modify this risk?

The distributions of (CTG)_n copy numbers among patients with congenital DM and among those with DM of later onset have been evaluated in two studies.^{13,14} The data from these studies are summarised in table 1B and indicate that the distributions of (CTG)_n copy numbers in these two groups of subjects differ. A mutation analysis of affected mother/affected child pairs¹⁴ has indicated a second feature

Table 1 Bayesian estimation of the probability that a baby carrying the DM mutation will have congenital myotonic dystrophy (CMD).

The prior probability is estimated on the basis of the maternal phenotype.⁴

A Maternal phenotype during pregnancy	Proportion of children with the mutation who will have CMD	
Normal	0.01*	(n = 32)
Multisystem disease	0.80	(n = 148)
Previous child with CMD	1.0	(n = 7)

The conditional probabilities are determined by the degree of fetal (CTG)_n amplification,^{13,14,24} and the increase in the degree of (CTG)_n amplification in the fetus compared with the mother.¹⁴

Table 1B lists the proportions of subjects with CMD or non-CMD whose mutations fall within the specified range.

B Degree of (CTG) _n amplification†: Approximate (CTG) _n copy number:	0 to <1.5 kb 50-499	1.5 to <3.0 kb 500-999	3.0 to <4.5 kb 1000-1499	Greater than 4.5 kb 1500 or more
Proportion of subjects with CMD‡ (n = 46)	0.01*	0.11	0.33	0.55
Non-CMD (n = 194)	0.47	0.38	0.14	0.01

Table 1C lists the proportion of subjects with CMD or non-CMD in whom the degree of (CTG)_n amplification was greater than that of their mothers.¹⁴

C	Degree of fetal amplification same as mother's	Degree of fetal amplification greater than mother's§
Proportion of subjects with CMD	0.01*	0.99 (n = 22)
Non-CMD	0.21	0.79 (n = 42)

* This figure is a conservative estimate of the true proportion; no instances were observed in the studies.

† The degree of amplification is based on the band shift noted on Southern blot analysis of genomic DNA.^{13,14,24} The range '0 to <1.5 kb' incorporates classes E₀ and E₁.^{14,24}

‡ The children with early childhood disease¹³ are excluded from this analysis.

§ An increase in the degree of (CTG)_n amplification is defined as the fetus having a higher molecular weight band than that of the mother identified on a Southern blot.¹⁴

that may be useful in estimating the risk of a baby having congenital DM. Each of the 22 congenitally affected children had a greater degree of (CTG)_n amplification (determined by Southern blot) than their mothers. Of the 42 non-congenitally affected children, 79% had (CTG)_n amplification to a greater degree than their mothers.

If these molecular distinctions are confirmed, the figures derived from the molecular studies could be combined with the empirical risks based on clinical data by using Bayes's theorem. An application of the clinical and molecular data is shown in tables 1 and 2. It

can be seen that using both sources of data can make a substantial difference to the calculated risks. The conditional probabilities listed in table 1 must be regarded as tentative as the number of cases studied remains small. Determining the (CTG)_n copy number of the fetus is relatively straightforward: as shown in this issue of the Journal, the (CTG)_n repeat can be detected in chorionic tissue, and shows relatively little of the somatic mosaicism evident in blood samples from older patients.²⁵ It may well be that a more refined algorithm incorporating both the clinical status of the mother and her previous children and the (CTG)_n copy

Table 2 Estimation of risks of congenital myotonic dystrophy (CMD) combining clinical and molecular data.

A woman has no previous children, multisystem DM, and (CTG)_n amplification of 2.0 kb; her fetus has (CTG)_n amplification of >4.5 kb. What is the probability that the baby will have CMD?

	Child will have CMD	Child will not have CMD
Prior probability of child having CMD	0.80	0.20
Conditional probability of fetal (CTG) _n amplification being >4.5 kb	0.55	0.01
Conditional probability of fetal (CTG) _n amplification being greater than mother's	0.99	0.79
Joint probability	0.44	0.002
Posterior probability	0.99	0.01

Therefore, the risk of child having congenital myotonic dystrophy is approximately 99%.

A woman has no previous children, multisystem DM, and (CTG)_n amplification of 2.0 kb; her fetus also has (CTG)_n amplification of 2.0 kb. What is the probability that the baby will have CMD?

	Child will have CMD	Child will not have CMD
Prior probability of child having CMD	0.80	0.20
Conditional probability of fetal (CTG) _n amplification being 2.0 kb	0.11	0.38
Conditional probability of fetal (CTG) _n amplification being the same as mother's	0.01	0.21
Joint probability	0.0009	0.016
Posterior probability	0.05	0.95

Therefore, the risk of the child having congenital myotonic dystrophy is approximately 5%.

number in the mother and fetus (ideally determined by PCR) could be developed.

Population genetics of the DM mutation

The consistent association of increased (CTG)_n copy number and DM in different populations is not surprising given the high degree of linkage disequilibrium between the DM mutation and nearby polymorphic loci.^{5,9,21,26} The disequilibrium data suggest that the majority of mutant alleles were derived from one ancestral mutation,²⁷ thus confirming the clinical impression that new mutations are very rare.¹ (Conversely, in this issue Davies *et al*¹² show that the distribution of (CTG)_n copy numbers in *normal* Japanese subjects differs from the distribution found among Caucasians.)

The difficulty lies in reconciling the molecular instability of the (CTG)_n repeat with the long history of the mutation. The mutant allele must remain quiescent and relatively stable for many generations. At some point the (CTG)_n repeat becomes unstable, and a molecular and phenotypic cascade ensues in successive generations. The cascade evolves very rapidly: the change from asymptomatic carrier with minimal features (and normal genetic fitness) to catastrophic congenital DM occurs in just a few generations. A similar phenomenon has been documented in the fragile X syndrome. Linkage disequilibrium between the fragile X (CCG)_n repeat and adjacent polymorphisms indicates that there have been relatively few primordial mutations.²⁸ The 'high risk' haplotype is widely distributed within the normal population and is associated with a (CCG)_n copy number at the upper end of the normal range; a similar evaluation of haplotypes and (CTG)_n copy number in a non-myotonic population has yet to be reported. A parsimonious explanation is that both triplet repeats mutate in a series of steps, with each increase in copy number increasing the further mutability of the repeat by some orders of magnitude.^{29,30} The counselling implications of this are profound. At what point does the mutation constitute a sufficient risk to a person to warrant population screening, clinical evaluation, counselling, or prenatal diagnosis?

Definition of molecular pathogenesis

The association of an increase in (CTG)_n copy number and DM is very strong, and provides presumptive evidence that the (CTG)_n repeat is the mutation. The repeat is in a non-coding region of DNA, and could regulate the expression of a gene upstream or downstream from the repeat. The myotonin protein kinase gene is an attractive candidate for the DM gene itself, and expression of various isoforms of the gene may account, in part, for the protean manifestations of the disease.³¹ A second gene lies immediately 5' to the myotonin protein kinase gene,³¹ and this gene may also be involved in the disorder. However, it remains to be shown that expression of either gene is

abnormal in DM, that such abnormal expression is necessary and sufficient to cause the disease, and that reduction of (CTG)_n copy number to the normal range is associated with correction of any yet to be identified abnormal cellular biochemistry. It is not known what factors account for the phenotypic variation noted at any given (CTG)_n copy number. Nor is it known why congenital DM is exclusively a consequence of maternal transmission.

Conclusion

The clinical utility of molecular analysis of the (CTG)_n repeat in the management of DM may be summarised as follows.

First, molecular analysis is an accurate and efficient means of identifying the mutation and confirming a clinical diagnosis. It is the investigation of choice in the diagnosis of subjects with equivocal signs, defining carriers who are asymptomatic, and determining whether a fetus has inherited the mutation. In view of the general relationship between (CTG)_n copy number and severity of disease, clinically significant mutations can probably be identified using Southern blot analysis. However, this approach may fail to detect mutations where the (CTG)_n copy number is less than 100 to 150, and in some instances analysis by PCR is essential. Conversely, some mutations may fail to be amplified by PCR.^{10-12,20} The diagnostic laboratory must be in a position to perform both procedures.

Second, in the current state of knowledge, molecular analysis generally cannot provide a precise indication of the severity of the disease. The one exception is that the combination of careful clinical evaluation and molecular studies can provide a useful estimate of the risk that a fetus will be born with congenital DM. A clinically useful guide to the genotype/phenotype relationship in DM will require careful integration of clinical, molecular, and biochemical studies in a large number of patients. It is essential that the method of reporting clinical and molecular data be uniform and consistent.

It is to be hoped that these scientific developments will encourage clinicians to take a greater interest in the overall management of DM families, and generally increase the awareness of the condition. The recent formation of a specific DM support group in the UK is encouraging in this regard*. Many affected subjects are not offered long term follow up to screen for disease complications such as cardiac arrhythmias³² and have not been warned about the possibility of anaesthetic complications.¹ It is not unusual to diagnose DM in an older woman and find that she has a child with mental retardation owing to unrecognised congenital DM. A lack of awareness of the bowel complications in congenital DM has led to the misdiagnosis of child abuse in some cases.³³

DM has joined the list of common conditions for which direct molecular evaluation

*Myotonic Dystrophy Support Group, C/o 175A Carlton Hill, Carlton, Nottingham NG4 1GZ, UK.

of the underlying mutation has replaced tests that may be more invasive, more expensive, and less accurate. Health care providers need to be cognisant of these developments, and to be aware of the need to address the inadequate staffing levels of most diagnostic DNA laboratories.

GKS and KED are supported by the Nuffield Foundation, the Muscular Dystrophy Group of Great Britain & Northern Ireland, and The Medical Research Council.

GRAEME K SUTHERS
SUSAN M HUSON
KAY E DAVIES

*Molecular Genetics Group,
Institute of Molecular Medicine,
John Radcliffe Hospital,
Headington, Oxford OX3 9DU, and
Department of Medical Genetics,
Churchill Hospital,
Oxford.*

Note added in proof

Biancalana *et al*³⁴ have now shown that the amplified androgen receptor repeat in SBMA families displays a low level of instability in both meiosis and mitosis, which is similar to that documented in fragile X and DM premutations.

- Harper PS. *Myotonic dystrophy*. 2nd ed. London: Saunders, 1989.
- Harper PS, Harley HG, Reardon W, Shaw DJ. Anticipation in myotonic dystrophy: new light on an old problem. *Am J Hum Genet* 1992;51:10-16.
- Brunner HG, Smeets HJM, Nillesen, W, *et al*. Myotonic dystrophy: predictive value of normal results on clinical examination. *Brain* 1991;114:2303-11.
- Koch MC, Grimm T, Harley HG, Harper PS. Genetic risks for children of women with myotonic dystrophy. *Am J Hum Genet* 1991;48:1084-91.
- Harley HG, Brook JD, Rundle SA, *et al*. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* 1992;355:545-6.
- Buxton J, Shelbourne P, Davies J, *et al*. Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* 1992;355:547-8.
- Aslanidis C, Jansen G, Amemiya C, *et al*. Cloning of the essential myotonic dystrophy region and mapping of the putative defect. *Nature* 1992;355:548-51.
- Fu YH, Pizzuti A, Fenwick RG, *et al*. An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* 1992;255:1256-9.
- Mahadevan M, Tsilfidis C, Sabourin L, *et al*. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* 1992;255:1253-5.
- 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.
- Brunner HG, Nillesen W, van Oost BA, *et al*. Presymptomatic diagnosis of myotonic dystrophy. *J Med Genet* 1992;29:780-4.
- Davies J, Yamagata H, Shelbourne P, *et al*. Comparison of the myotonic dystrophy associated CTG repeat in European and Japanese populations. *J Med Genet* 1992;29:766-9.
- Harley HG, Rundle SA, Reardon W, *et al*. Unstable DNA sequence in myotonic dystrophy. *Lancet* 1992;339:1125-8.
- Tsilfidis C, MacKenzie AE, Mettler G, Barcelo J, Korneluk RG. Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. *Nature Genet* 1992;1:192-5.
- Sutherland GR, Haan EA, Kremer E, *et al*. Hereditary unstable DNA: a new explanation for some old genetic questions? *Lancet* 1991;338:289-92.
- Nakahori Y, Knight SLJ, Holland J, *et al*. Molecular heterogeneity of the fragile X syndrome. *Nucleic Acids Res* 1991;19:4355-9.
- Fu YH, Kuhl DPA, Pizzuti A, *et al*. Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 1991;67:1047-58.
- Yu S, Mulley J, Loesch D, *et al*. Fragile-X syndrome: unique genetics of the heritable unstable element. *Am J Hum Genet* 1992;50:968-80.
- La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fishbeck KH. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 1991;352:77-9.
- Caskey CT, Pizzuti A, Fu YH, Fenwick RG, Nelson DL. Triplet repeat mutations in human disease. *Science* 1992;256:784-9.
- Yamagata H, Miki T, Ogihara T, *et al*. Expansion of unstable DNA region in Japanese myotonic dystrophy patients. *Lancet* 1992;339:692.
- Melchionda S, Cobo A, Gennarelli M, *et al*. Expansion of the myotonic dystrophy gene in Italian and Spanish patients. *J Med Genet* 1992;29:789-90.
- Reardon W, Harley HG, Brook JD, *et al*. Minimal expression of myotonic dystrophy: a clinical and molecular analysis. *J Med Genet* 1992;29:770-3.
- Hunter A, Tsilfidis C, Mettler G, *et al*. The correlation of age of onset with CTG trinucleotide repeat amplification in myotonic dystrophy. *J Med Genet* 1992;29:774-9.
- Myring J, Meredith AL, Harley HG, *et al*. Specific molecular prenatal diagnosis for the CTG mutation in myotonic dystrophy. *J Med Genet* 1992;29:785-8.
- Harley HG, Brook JD, Floyd J, *et al*. Detection of linkage disequilibrium between the myotonic dystrophy locus and a new polymorphic DNA marker. *Am J Hum Genet* 1991;49:68-75.
- Ashizawa T, Epstein HF. Ethnic distribution of myotonic dystrophy gene. *Lancet* 1991;338:642-3.
- Richards RI, Holman K, Friend K, *et al*. Evidence of founder chromosomes in fragile X syndrome. *Nature Genet* 1992;1:257-60.
- Morton NE, Macpherson JN. Population genetics of the fragile-X syndrome: multiallelic model for the FMR1 locus. *Proc Natl Acad Sci USA* 1992;89:4215-7.
- Chakravarti A. Fragile X founder effect? *Nature Genet* 1992;1:237-8.
- Jansen G, Mahadevan M, Amemiya C, *et al*. Characterization of the myotonic dystrophy region predicts multiple protein isoform-encoding mRNAs. *Nature Genet* 1992;1:261-6.
- Anonymous. The heart in myotonic dystrophy. *Lancet* 1992;339:528-9.
- Reardon W, Hughes HE, Green SH, Woolley VL, Harper PS. Anal abnormalities in childhood myotonic dystrophy - a possible source of confusion in child sexual abuse. *Arch Dis Child* 1992;67:527-8.
- Biancalana V, Serville F, Pommier J, Julien J, Hanauer A, Mandel JL. Moderate instability of the trinucleotide repeat in spino-bulbar muscular atrophy. *Hum Mol Genet* 1992;1:255-8.