ORIGINAL ARTICLES

Comparison of the myotonic dystrophy associated CTG repeat in European and Japanese populations

June Davies, Hidehisa Yamagata, Peggy Shelbourne, Jessica Buxton, Toshio Ogiwara, Pekka Nokelainen, Masanori Nakagawa, Robert Williamson, Keith Johnson, Tetsuro Miki

Department of
Anatomy, Charing
Cross and
Westminster Medical
School, St Dunstan's
Road, London W6
8RP.

J Davies
P Shelbourne
J Buxton*
K Johnson

Department of
Geriatric Medicine,
Osaka University
Medical School,
Japan.

H Yamagata
T Ogiwara
T Miki

Laboratory of Medical
Genetics, National
Public Health
Institute, Helsinki,
Finland.
P Nokelainen

Third Department of
Internal Medicine,
Faculty of Medicine,
Kagoshima
University, Japan.
M Nakagawa

Department of
Biochemistry and
Molecular Genetics,
St Mary's Hospital
Medical School,
Imperial College,
London.
R Williamson

*Present address:
Mothercare Unit of
Paediatrics, Institute of
Child Health, London
WC1N 1EH.

Correspondence to
Dr Johnson.

Received 17 June 1992.
Revised version accepted
13 July 1992.

Abstract

Gene amplification using polymerase chain reaction (PCR) was carried out on DNA samples from a total of 92 normal subjects and 52 subjects with myotonic dystrophy (DM) from European and Japanese populations, to determine the copy number of the CTG repeat associated with DM for each group. In the two populations, the number of repeats on normal chromosomes only were compared, as CTG copy number on DM chromosomes was difficult to determine by PCR alone. In this study, normal chromosomes were found which had as many as 35 copies of the repeat, which is larger than the normal range reported previously but still does not overlap with the repeat number associated with DM pathology, which is at least 50 copies. Using data from normal chromosomes from unrelated subjects, the frequencies of five, 11, and 13 copies of the CTG repeat were found to be significantly different between the two populations, with five and 11 copies more commonly seen in the European population and 13 copies in the Japanese population. This difference may be the result of natural divergence of the normal chromosomes between the population groups.

(*J Med Genet* 1992;29:766-9)

Myotonic dystrophy (DM) is the most common form of adult muscular dystrophy, with an average incidence of 1 in 8000 in the European population and a lower incidence of 1 in 20 000 in Japan.¹² DM is characterised by myotonia with muscle wasting and a wide range of accompanying symptoms including cataracts, intellectual impairment, and conduction defects in the heart.¹ The age of onset and severity of the disease show extreme variation both between and within families.

The genetic defect underlying DM was unknown until recently, when it was reported that a fragment of DNA is vastly expanded in DM patients.³⁻⁵ This expansion is the result of an increased number of a trinucleotide CTG

repeat at the 3' end of a gene encoding a member of the protein kinase family.⁶⁻⁸

The repeat copy number has been found to be highly variable in the normal population, with a reported size range between 5 and 30 copies.⁶⁻⁸ DM patients are found to have one allele in the normal size range and a larger DM specific allele; mildly affected or presymptomatic subjects have at least 50 copies of the repeat and severely affected patients have alleles containing up to several thousand copies of CTG. The extent of the expansion of the CTG repeat correlates well with age of onset and increased severity of symptoms within families: the larger the number of CTGs, the more severe the effects of the disease, but at least 50 copies are necessary to produce detectable clinical symptoms.⁶⁻⁸ The number of repeated sequences can change in succeeding generations, and an increase in repeat number indicates why children born to minimally affected parents can show more severe symptoms. This explains the phenomenon of genetic anticipation seen in myotonic dystrophy families.^{9,10}

Polymerase chain reaction (PCR) using synthetic oligonucleotides flanking the CTG repeat allows the specific amplification of the region containing the repeat. The copy number of the repeat is very variable, and has been shown to decrease as well as increase.¹¹ It is important for greater accuracy in diagnosis to establish the normal range of alleles. This study directly compares the number of copies of CTG on non-DM chromosomes in European and Japanese populations.

Methods

DNA was prepared from peripheral blood or immortalised lymphoblastoid cell lines.¹² A total of 144 unrelated subjects was studied, 37 normal and 32 DM subjects from Japan and 55 normal and 20 DM subjects from Europe including British, German, Belgian, Swedish, and Finnish families. The largest subgroup in the European population originated from Finland with 16 normal and 11 DM subjects. Statistical values were calculated using the χ^2 test in 2×2 tables with Yates's correction. The

Figure 1 Frequency distribution of the DM associated CTG repeat in European and Japanese populations. The European population contains families from Britain, Belgium, Germany, Sweden, and Finland, with the Finnish being the largest subgroup. No statistical difference in allele frequency was found between the Finnish subgroup and the rest of the European population.

polymorphic content of the CTG repeat was calculated using $PIC = 1 - \sum_{i=1}^n P_i^2$, where P_i is the frequency of allele i .¹³

PCR was carried out using 20 ng of DNA, 50 pmol of each primer 101 and 102,⁶ 200 µmol/l each of dATP, dTTP, dCTP, and dGTP (Pharmacia), 1 unit of *Tth* polymerase enzyme (Hybaid) in a final volume of 25 µl overlaid with two drops of mineral oil (Sigma). Radiolabel was added to the PCR samples in two ways: (1) end labelling of primer 101 using [γ -³²P]ATP (Amersham); 50 pmol of primer 101 was incubated at room temperature for 30 minutes with 2 units of T4 polynucleotide kinase in a 20 µl reaction containing 2 µl of 10 × kinase buffer (final concentration 50 mmol/l Tris-HCl pH 7.5, 10 mmol/l magnesium chloride, 5 mmol/l dithiothreitol, 0.1 mmol/l spermidine), and 1 µl [γ -³²P]ATP. 0.5 pmol of radiolabelled primer 101 was added to each PCR sample. (2) Direct incorporation of [α -³²P]dCTP (Amersham) by addition of 20 to 30 nCi [α -³²P] dCTP to each PCR sample.

PCR running conditions were as follows: initial denaturation at 94°C for five minutes, then 35 cycles of 94°C for 1.5 minutes, 62°C for one minute, and 72°C for two minutes. Five microlitres of the PCR product was first checked on a 2% agarose gel (containing equal amounts of Nusieve and regular agarose), run for one hour in 1 × TBE buffer. Then 3 µl of the PCR product was denatured with 2 µl

χ^2 values for the most common CTG alleles found in the European and Japanese populations. Comparisons were made separately for each allele using 2 × 2 tables with Yates's correction.

Copies of CTG	No of European chromosomes	No of Japanese chromosomes	χ^2	p
5	44	21	5.08	0.02 < p < 0.03
11	27	9	5.89	0.01 < p < 0.02
12	18	25	3.09	0.05 < p < 0.1
13	9	29	16.56	p < 0.001
14	9	1	4.66	0.05 < p < 0.1
Others	23	21		
Total	130	106		

formamide dye (USB Sequenase version 2.0 sequencing kit, Cambridge Bioscience) and run on a 6% denaturing polyacrylamide gel. The polyacrylamide gel was dried and exposed to Hyperfilm-MP (Amersham) for one to two days.

Results

Heterozygote frequencies in the normal population studied were 73% (European) and 81% (Japanese). The CTG repeat is very informative and has a PIC value of 0.83 based on the frequency of 19 alleles seen in this study. The CTG repeat alleles showed mendelian inheritance through families, and in 146 meioses studied no new mutations were seen on normal chromosomes but the mutation rate on the DM chromosomes was 100%.

The frequencies of the CTG repeat of normal alleles from the European and Japanese populations are shown in fig 1. The distribution of alleles is indicated, with a peak repeat number of five, and then a second peak between 11 and 14. Very few repeats were found between five and 11; only a single example of an eight copy allele (European) and four of 10 copies (three European and one Japanese) were found.

The modal number of copies is five in the European population but 13 in the Japanese. The results of comparing the most common alleles from unrelated subjects from the European and Japanese populations, using the χ^2 test in 2 × 2 tables with Yates's correction, are shown in the table. Linkage disequilibrium between the CTG repeat size and flanking markers on normal chromosomes was looked for: the sample size for the European population was too small to determine if disequilibrium was present, but preliminary data from the Japanese population indicate that disequilibrium exists between 12 and 13 copies of CTG and the probe p37.1. No disequilibrium was found with D19S63.

In looking at the normal range of CTG copies (including the non-DM chromosome from affected subjects) four different persons were seen with more than 30 copies of CTG, 32 and 35 copies on two Japanese chromosomes and 33 and 35 on two European chromosomes (fig 2). In pedigree c in fig 2A, a 35 copy allele is inherited in a mendelian fashion by an affected child (lane 7) on the non-DM chromosome as a normal allele from the unaffected parent (lane 9). In pedigree b in fig 2A, the reverse occurs where 33 copies of CTG are inherited as a normal allele by an unaffected child from the affected parent. These results give a normal range of five to 35 copies for the CTG repeat in these populations.

The DM allele is generally not amplified under the conditions used and the smallest allele sizes for detectable clinical symptoms were only seen in three cases. Under the PCR conditions used, two alleles were seen in three mildly affected subjects and in each case the second (DM) allele was visualised as a faint diffuse band (< 400 base pairs in size), which corresponded to repeat copy numbers of approximately 65, 78, and 86 (data not shown).

Figure 2 (A) Two Finnish pedigrees showing mendelian inheritance of 33 and 35 copies of the CTG repeat. The numbers on pedigrees b and c correspond to the numbers on the lanes of the autoradiograph in a. Lane 10 is a control sample containing 20 copies of CTG. Lanes 3 and 5 contain 33 copies of CTG, and lanes 7 and 9 contain 35 copies of CTG. (B) A Japanese pedigree showing mendelian inheritance of 35 copies of the CTG repeat. The numbers on pedigree b correspond to the numbers on the lanes of the autoradiograph in a. Lanes 1 and 2 are control samples containing 12, 23, and 5 copies of CTG respectively. The five copies of CTG seen in lane 8 in this gel is the result of contamination by PCR carryover and was not detected in other reactions.

Discussion

DM is caused by a mutation that increases the size of a genomic fragment by the amplification of a trinucleotide CTG repeat. This is similar to the trinucleotide repeat amplifications which cause mental retardation associated with fragile X (CGG) and X linked spinal and bulbar muscular atrophy (CAG).¹⁴⁻¹⁸ In this study we directly compared the number of copies of CTG on normal chromosomes for the European and Japanese populations by PCR. The results shown in fig 2 indicate a normal range of five to 35 copies for the CTG repeat in these populations, which suggests that the normal range of alleles is greater than previously reported but still does not overlap with the > 50 copies associated with DM pathology.⁶⁻⁸ In comparison, fragile X, which also shows anticipation and is associated with a trinucleotide repeat, has a normal range of between six and 54 copies of a CGG repeat and a premutation stage with 52 to over 200 copies of CGG.^{14,16,19} The alternative explanation, that these subjects with > 30 copies of the CTG repeat could possess a premutation, indicating predisposition to DM, is unlikely, as the 33 and 35 copy alleles were found on normal chromosomes in both normal and DM subjects. Further study of normal subjects is needed to determine the normal range of the CTG repeat length and whether, like fragile X, the DM trinucleotide repeat could have a premutation stage.

The frequencies of the normal alleles containing 5, 11, and 13 copies of CTG show a significant difference between these two populations (table 1), with a greater frequency of alleles 5 and 11 on European chromosomes and

allele 13 on Japanese chromosomes. The mutation originally detected by cDNA-25 as an expanded fragment has also been found in the Japanese population, confirming that the DM locus is not genetically heterogeneous.^{4,20} The cDNA-25 probe detects an *EcoRI* polymorphism with allele sizes of 9.8 kb and 8.6 kb, with no significant difference in the frequencies of the alleles between Japanese and European populations. Linkage disequilibrium data for the cDNA-25 *EcoRI* polymorphism and other polymorphic markers in the area with DM indicate that the DM mutation may have arisen as a common ancestral event.^{5,20-23} This seems to be borne out by the report of Ashizawa and Epstein²⁴ that the mutation responsible for DM occurred before the migration of North Eurasian groups from Africa to Europe and Japan.²⁴ Edwards *et al*²⁵ state that higher mutation rates are associated with a greater number of core repeats and this correlates well with the high mutation rate (100%) seen for > 50 copies of CTG on DM chromosomes. The low mutation rate for five to 35 copies of CTG on normal chromosomes could account, by natural divergence and genetic drift, for the differences in the frequency of CTG copy number seen between the Japanese and European population groups.

The DM mutation represents the third example of a disease caused by the amplification of a trinucleotide sequence within a gene, suggesting a common genetic mechanism which may be found to be the cause of many more genetic disorders. Since very few alleles are found between five and 11 copies of the CTG, the initial amplification event has been suggested to be a doubling or tripling of the

number of repeats with subsequent slippage,⁶ but data on the mechanism of sequence amplification during meiosis and mitosis are not yet available for trinucleotide sequences, such as those involved in DM.

PCR on DNA from most cases of DM patients gives only one band. This is because in severely affected patients the unstable fragment has increased by anything up to 5 kb, and the PCR reaction undergoes allele specific amplification where short stretches of DNA are amplified more efficiently in comparison with longer fragments. Most affected subjects therefore only show the PCR product from the unaffected chromosome. This could be a problem in using PCR alone for diagnostic purposes, as it is difficult to distinguish affected from normal subjects homozygous for a particular CTG copy number. The range of amplification sizes owing to mitotic instability of the DM chromosome will also make it more difficult to visualise any product that is made. This somatic heterogeneity has been seen as a diffuse range of fragments by Southern blotting of digested lymphoblastoid DNA and by PCR, where it becomes difficult to estimate the exact CTG copy number on the affected DM chromosome.^{4,5,11} Recent work on sequencing SBMA mutations and the detection of the full mutation for fragile X could now be applied to DM to overcome these problems.^{26,27}

In conclusion, the range of normal alleles has been extended from that previously reported, which aids in delineating the range of normal compared to DM alleles. Population differences have been found in the frequency of normal alleles in European and Japanese population groups. This can be explained by population divergence of the normal alleles, as the same mutation for DM has been found in both groups. At the moment PCR alone has a limited use as an informative diagnostic test and for the most accurate prediction Southern blots with *Bam*HI and *Bgl*II digests should be used for the initial screening¹¹ (P Shelbourne, in preparation).

Research at Charing Cross is supported by the MDG through grant number RA3/257/1, the MDA/Piton Foundation through a grant under its Task Force in Genetics and the University of London Central Research Fund. JB was on a SERC studentship. Research in Japan is supported by grant in aid for scientific research in priority areas from the Ministry of Education, Science, and Culture and grants from the Ministry of Health and Welfare, Japan. We would like to thank Clare Jones and Tracey van Tongeren for reading this manuscript, Terry Bull for photographic assistance, and Dr Ken MacCrae for statistical advice.

- 1 Harper PS. *Myotonic dystrophy*. 2nd ed. Philadelphia: Saunders, 1989.
- 2 Osame M, Furusho T. Genetic epidemiology of myotonic dystrophy in Kagoshima and Okinawa districts in Japan. *Clin Neurol* 1983;23:1067-71.
- 3 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.
- 4 Buxton JL, Shelbourne PF, Davies JL, et al. Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* 1992;355:547-8.
- 5 Harley HG, Brook JD, Rundle SA, et al. Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* 1992;355:545-7.
- 6 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.
- 7 Fu YH, Pizzuti A, Fenwick RG, et al. An unstable triplet repeat in a gene related to myotonic muscular dystrophy, myotonin protein kinase. *Science* 1992;255:1256-8.
- 8 Mahadevan M, Tsifildis C, Sabourin L, et al. Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of a candidate gene. *Science* 1992;255:1253-5.
- 9 Howeler CJ, Busch HFM, Geraedts JPM, Niermeijer MF, Staal A. Anticipation in myotonic dystrophy: fact or fiction. *Brain* 1989;112:779-97.
- 10 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.
- 11 Shelbourne P, Davies J, Buxton J, et al. Direct diagnosis of ambiguous cases of myotonic dystrophy using the disease-specific DNA marker. *N Engl J Med* (submitted).
- 12 Nimmo E, Johnson K. RFLPs: uses and methods of detection. In: Dale JW, Sanders PG, eds. *Methods in gene technology*. London: JAI Press, 1991:267-83.
- 13 Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* 1980;32:314-31.
- 14 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;67:1-20.
- 15 Verkerk AJ, Pieretti M, Sutcliffe JS, et al. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* 1991;66:905-14.
- 16 Yu S, Pritchard M, Kremer E, et al. Fragile X genotype characterised by an unstable region of DNA. *Science* 1991;252:1179-81.
- 17 La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischback KH. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 1991;352:77-9.
- 18 Rousseau F, Heitz D, Biancalana V, et al. Direct diagnosis by DNA analysis of the fragile X syndrome of mental retardation. *N Engl J Med* 1991;325:1673-81.
- 19 Oberlé I, Rousseau F, Heitz D, et al. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science* 1991;262:1097-102.
- 20 Yamagata H, Miki T, Ogiwara T, et al. Expansion of unstable DNA region in Japanese myotonic dystrophy patients. *Lancet* 1992;339:692.
- 21 Harley HG, Walsh KV, Rundle S, et al. Localisation of the myotonic dystrophy locus to 19q13.2-19q13.3 and its relationship to twelve polymorphic loci on 19q. *Hum Genet* 1991;87:73-80.
- 22 Harley H, Brook JD, Rundle S, et al. Detection of linkage disequilibrium between the myotonic dystrophy locus and four DNA polymorphisms in European and other populations. *Am J Hum Genet* 1991;49 (suppl):1916A.
- 23 Nokelainen P, Alanen-Kurki L, Winquist R, et al. Linkage disequilibrium detected between dystrophin myotonia and APOC2 locus in the Finnish population. *Hum Genet* 1990;85:541-5.
- 24 Ashizawa T, Epstein H. Ethnic distribution of myotonic dystrophy gene. *Lancet* 1991;338:642.
- 25 Edwards A, Hammond H, Jin L, Caskey T, Chakraborty R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* 1992;12:241-53.
- 26 Pergolizzi RG, Erster SH, Goonewardena P, Brown WT. Detection of full fragile X mutation. *Lancet* 1992;339:271-2.
- 27 Yamamoto Y, Kawai H, Nakahara K, et al. A novel primer extension method to detect the number of CAG repeats in the androgen receptor gene in families with X-linked spinal and bulbar muscular atrophy. *Biochem Biophys Res Commun* 1992;182:2507-13.