Absence of myotonic dystrophy in southern African Negroids is associated with a significantly lower number of CTG trinucleotide repeats

Andrea Goldman, Michele Ramsay, Trefor Jenkins

Abstract
Myotonic dystrophy (DM) is associated with an increased number of CTG repeats in the 3’ untranslated region of the myotonin gene. Because DM has not been observed in southern African Negroids, a study of the CTG repeat polymorphism in this population was undertaken. A total of 210 unrelated subjects was studied by PCR analysis of the trinucleotide repeat in the DM gene and the size and distribution of the CTG repeat were determined. The alleles ranged in length from five to 22 repeats. A previously undescribed BgII polymorphism was found which could lead to erroneous diagnosis of DM in people from this population. South African Negroids were found to have significantly fewer large repeat lengths than do white and Japanese populations. It is suggested that the occurrence of fewer large CTG repeats in the normal range may, in part, explain the absence of DM in southern African Negroids.

(J Med Genet 1994;31:37-40)

Myotonic dystrophy (DM) is an autosomal dominant disorder that occurs with a prevalence of 1:8000 worldwide. The classical clinical features include myotonia, progressive muscle weakness and wasting, and many associated symptoms, including cataracts, intellectual impairment, and frontal balding. A hallmark of the disease is the extreme variation in severity and age of onset, both between and within families. Although the biochemical defect underlying this disorder is still unknown, the genetic defect has recently been shown to result from an unstable CTG repeat DNA sequence in the 3’ non-coding region of the myotonin protein kinase gene, located on chromosome 19q13.3. When amplified, this trinucleotide repeat is responsible for DNA instability and molecular pathology. In normal subjects, the CTG repeat is highly polymorphic, with the copy number ranging from five to 38. In mildly affected myotonic dystrophy patients, the copy number is greater than 50, and has been found to be as high as 2000 in severely affected patients. The expansion in size of the CTG repeat correlates well with disease severity and age of onset within families.

Although diagnosis of DM can be difficult owing to the wide range and severity of symptoms, the disorder has been described from many countries worldwide, including Japan, China, India, South Africa, Nigeria, and also in black American families. No single case has, however, been reported in an indigenous Negroid or Khoisan person from southern Africa, despite a recent survey representing a population of more than 30 million. This survey, which also included other African countries, concluded that DM has a very low prevalence among ethnic Africans, especially in central and southern Africa.

Since DM has not been described in Negroids, a study was undertaken to verify the presence of the CTG repeat polymorphism in a Negroid population with minimal white admixture. The distribution of the CTG repeat in normal Negroid subjects was determined in an attempt to establish whether CTG repeat length correlated with the absence of DM in this population. It was hypothesised that smaller numbers of CTG repeats within the normal range would occur in the Negroid population and that these alleles may be more stable than larger alleles in the normal range.

Subjects and methods
A total of 210 unrelated southern African Bantu-speaking Negroids was studied. They were drawn from the various South African chiefdoms and were, in the main, male blood donors or staff of our Institute. Genomic DNA was extracted from peripheral blood using either a phenol-chloroform extraction or a simple salting out procedure. The polymerase chain reaction (PCR) with primers 101 and 102 was used to amplify the region containing the CTG repeat. The PCR reaction was performed in a 25 µl reaction volume containing 100 ng of genomic DNA, 10 mmol/l Tris-HCl, pH 8.8, 50 mmol/l KCl, 1.5 mmol/l MgCl₂, 0.1% Triton-X, 200 µmol/l of dATP, dCTP, and dGTP, and 2.5 µmol of dCTP, 50 ng of each primer, 20–30 ng of (α-32P)dCTP, and 2 U Taq DNA polymerase. A total of 35 cycles of amplification (94°C for one minute, 66°C for one minute, 72°C for one minute) was carried out using a Perkin Elmer Cetus DNA Thermal Cycler. The denatured PCR products were run on 6% polyacrylamide gels. Samples with known CTG repeat numbers were run on every gel as controls and to facilitate the sizing of alleles. All samples found to have only one
Results
The variability in CTG repeat length was analysed for 420 chromosomes and the findings are shown in fig 1 and table 1. The genotypes of the subjects are not presented but are available to interested colleagues on request. A total of 14 alleles was found, ranging in size from five to 22 repeats. The most common allele had five repeats (27% of chromosomes) and the majority of alleles were in the range of 10 to 14 repeats (62%), with 12 repeats being most frequent (21%) in this group. No chromosome was observed with nine repeats and very few alleles were found with more than 14 repeats (5%). It is noteworthy that only three alleles were found to be longer than 19 repeats (0.7%). This system has a PIC value of 0.84 in the southern African Negroids.

Of the 210 subjects analysed using PCR alone, 36 appeared to be homozygous at the DM locus (16 subjects appeared to be 5/5, nine 12/12, six 13/13, four 11/11, and one 7/7). Southern blot analysis of BglII digested DNA probed with p581.4 confirmed homozgyosity in 32 subjects, but showed a larger (>3.4 kb) fragment size in four. Further analysis was carried out using BamHI digested DNA probed with p581.4, a system reported to have greater resolution in detecting variation of the CTG repeat than the BglII system. This failed to show the larger fragments seen with BglII, suggesting that the larger sized BglII fragments are the result of a polymorphism at a BglII restriction enzyme recognition site, rather than being caused by a (CTG)n expansion. This polymorphism has not been described previously.

The findings of this study were compared with those of Davies et al, who described the allele frequencies in non-DM subjects of European and Japanese populations. Comparisons for each allele were done using the χ2 test (table 1). Significant differences were found for alleles 10 and 11 when comparing the southern African Negroids and the European populations and the differences were significant for alleles 10, 13, and 14 when comparing the southern African Negroid and Japanese populations. Also, the χ2 analysis of the overall distribution showed that the differences are highly significant (Negroids v European, χ2[16]=56-0, p<0.0001; and Negroids v Japanese, χ2[16]=59-0, p<0.0001).

When alleles were divided into two groups, five to 14 repeats and greater than 14 repeats, Negroids appeared to have fewer alleles in the second group (0.05, SD 0.01) compared to Europeans (0.15, SD 0.03) and Japanese (0.19, SD 0.04) (inset in fig 2). These differences are highly significant (Negroids v European, χ2[16]=13.6, p=0.0002; Negroids v Japanese, χ2[16]=22.6, p<0.0001). If one were to analyse these data using the cut off point of Imbert et al (that is, the alleles were divided into two groups of five to 18 repeats and greater than 18 repeats), the differences between the three groups is even more significant (Negroids v European, χ2[16]=33.4, p<0.001; Negroids v Japanese, χ2[16]=19.6, p<0.001) (table 2).

Discussion
The DM mutation involves the expansion of an unstable trinucleotide repeat (CTG) in the 3' untranslated region of the myotonin gene. A similar mechanism of trinucleotide repeat expansion has been described in fragile X mental retardation syndrome (CGG), spinobulbar muscular atrophy (CAG), and more recently, Huntington's disease (CAG) and spinocerebellar ataxia type I. Myotonic dystrophy has never been described in the southern African Negroid population, so a study on the CTG trinucleotide repeat in this population was undertaken in an attempt to explain its apparent absence. The CTG repeat polymorphism is present and 14 alleles of different repeat lengths were observed. Similar findings have been reported in other populations, where alleles 5, 12, and 13 are most common. An unexpected but potentially important finding was the discovery of a BglII RFLP in

Table 1 χ2 comparisons of southern African Negroid (N), European (E), and Japanese (J) populations, with Yates's correction where necessary, for the common CTG repeat alleles, 5-14, and the pooled others

<table>
<thead>
<tr>
<th>No of CTG repeats</th>
<th>No of chromosomes</th>
<th>p value</th>
<th>N x E</th>
<th>N x J</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>112</td>
<td>44</td>
<td>0.033</td>
<td>0.147</td>
</tr>
<tr>
<td>10</td>
<td>45</td>
<td>3</td>
<td>0.005</td>
<td>0.040</td>
</tr>
<tr>
<td>11</td>
<td>47</td>
<td>27</td>
<td>0.005</td>
<td>0.040</td>
</tr>
<tr>
<td>12</td>
<td>87</td>
<td>18</td>
<td>0.082</td>
<td>0.519</td>
</tr>
<tr>
<td>13</td>
<td>50</td>
<td>9</td>
<td>0.109</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>14</td>
<td>33</td>
<td>9</td>
<td>0.028</td>
<td>0.009</td>
</tr>
<tr>
<td>Others</td>
<td>46</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>420</td>
<td>130</td>
<td>106</td>
<td></td>
</tr>
</tbody>
</table>

* Davies et al.14
the southern African Negroid population that has not been previously described. This RFLP may interfere with interpretation of results with the BglII/p5Bl.4 system and could lead to erroneous diagnosis of DM in unaffected persons. It is possible, however, that this polymorphism may be confined to African peoples, that is, it is Negro specific.

The allele distribution in this southern African Negroid population, using either 14 or 19 repeats as the cut off point, was significantly different from the distribution in two previously described populations, namely Europeans and Japanese. In the South Africans, there is a paucity of larger repeats. In an attempt to explain the possible implications of this finding, it is instructive to compare and contrast DM with the other disorders caused by a similar mutation mechanism. There are two important considerations: (1) the effect of large repeat numbers on the stability of the triplet repeat, and (2) the evidence of founder chromosomes in the syndromes.

The molecular mechanisms that determine the stability of triplet repeats during meiosis are unclear at present. Triplet repeats mutate in a series of steps, with each increase in copy number increasing the further mutability of the repeat by some order of magnitude. A certain ‘threshold’ number of repeats is required before replication fidelity is compromised and the gene function/product is disrupted. Caskey et al. suggested that the possible mechanism of triplet repeat expansion or instability may result from the difficulty of replication of GC rich sequences, thus causing inequality of DNA synthesis rate. This model predicts that longer alleles (maybe even those within the normal range) would be more prone to this mechanism, resulting in expansion.

Studies of fragile X families by Heitz et al. and Caskey et al. indicated that the size of the premutation is a major (and perhaps the only) determinant of the risk of transition from premutation to full mutation. There is thus a much greater risk of a large (≥ 90 repeats) premutation being transformed to a full mutation. To extrapolate from these studies in fragile X families, it is possible that, at the DM locus, the larger copy number within the normal range may have a greater chance of progressing to the mutation range of repeats and hence the manifestation of DM in the next generation. The recent findings of Imbert et al. also strongly suggest that the length of the CTG repeat in the DM gene is the main driving force for DM mutations. The authors proposed that the larger normal sized alleles with greater than 19 CTG repeats may have significant instability and thus may be the predisposing alleles for DM mutations.

Haplotype analysis of fragile X chromosomes showed that a few founder (ancestral) mutations are responsible for most fragile X cases. The ‘high risk’ haplotypes are widely distributed within the normal population and analysis of normal subjects with one of these haplotypes showed a high incidence of large CGG copy number at the upper end of the normal range. Similarly, investigations of linkage disequilibrium between Huntington’s disease (HD) and DNA markers have suggested that at least a third of HD chromosomes are ancestrally related. In two HD families,
in which a new mutation was thought to have occurred, the most common haplotype seen on HD chromosomes was found with a CAG copy number in the upper end of the normal range. The clinical appearance of HD in these two cases was associated with expansion of the CAG repeat. These findings suggest that there may be a pool of subjects with repeat numbers in the upper but normal range whose descendents are at an increased risk of subsequent mutation owing to the relationship between copy number and mutability.

We hypothesised that South African Negroids, a population in which DM has not, to our knowledge, been described, would have a lower frequency of the large CAG repeat alleles in the normal range than do populations in which the disease occurs at relatively high prevalence. This has been confirmed. The molecular mechanisms that determine the stability of DM alleles during inheritance are unclear at present. The findings in fragile X, the two HD families, and, more recently, the study of the DM locus in 149 subjects may suggest that, at the DM locus, the larger CAG copy number (alleles) within the normal range may be particularly susceptible to expansion into the mutation range. This may explain the absence of DM in South African Negroids. Expansion of the repeat lengths to 19 or more CTG repeats is in complete linkage disequilibrium with allele 1 (insertion) in the insertion/deletion polymorphism, thus suggesting the existence of a primordial predisposing mutation.

The CTG repeat lengths of 19 or more have been shown to be very rare in South African Negroids and it will be of interest to determine the frequencies of, and association with, the alleles (1 and 2) of the insertion/deletion polymorphism, as well as the frequencies of other closely linked polymorphic markers.

We thank Keith Johnson for the gift of the probe pSB1.4.