French myotonic dystrophy families show expansion of a CTG repeat in complete linkage disequilibrium with an intragenic 1 kb insertion

C Lavedan, H Hofmann-Radvanyi, C Boileau, C Bonaiti-Pellie, D Savoy, P Shelbourne, C Duros, J-P Rabes, I Dehaupas, S Luce, K Johnson, C Junien

Abstract

The molecular basis of myotonic dystrophy (DM) has been characterised. All DM mutations characterised to date appear as an unstable elongation of a fragment containing a tandem repeat of a CTG motif, which can be visualised in both EcoRI and BamHI digests. It has been shown that the fragment is polymorphic in the normal population. Another 1 kb insertion/deletion polymorphism located near the unstable CTG repeat region has been identified. The 1 kb insertion allele is present in all DM patients. These different polymorphic systems can be distinguished using cDNA25 and BamHI, because this enzyme cuts between the site of the 1 kb insertion and the CTG repeat. We thus haplotyped DM patients from 72 French families and clearly showed that all chromosomes (100%) with the DM mutation carried the 1 kb insertion as well. In addition to this association, we detected significant linkage disequilibrium between the DM locus and D19S53 for which allele frequencies were different from other European populations. Our results in the French DM population are thus in agreement with the hypothesis that the CTG expansion occurred on one or a few ancestral chromosomes carrying the large 1 kb insertion allele.

Myotonic dystrophy (DM) is an autosomal dominant, multisystemic disorder characterised by a highly variable clinical phenotype including myotonia, progressive muscle weakness, and cataracts. It is the most common form of adult muscular dystrophy. Myotonic dystrophy can occur as a severe congenital form only when transmitted by some carrier women. The DM mutation has been characterised as a CTG repeat amplification in the 3' untranslated region of the myotatonin protein kinase (MT-PK) gene. The size of this repeat sequence varies between affected sibs and increases through generations in parallel with increasing severity of the disease. Infants with severe congenital DM have been shown to have on average a greater amplification of the CTG repeat. The finding of the same mutation in many populations has suggested a common ancestry for the majority of cases. Accordingly, linkage disequilibrium has been detected between polymorphic DNA markers within the gene for apolipoprotein C2 (APOC2) and the DM locus in isolated populations, namely French Canadian and Finnish. In 1991, Harley et al reported the first linkage disequilibrium between DM and a tightly linked anonymous marker (D19S63) in a heterogeneous population; a similar result was later reported in Spanish and Japanese populations. This suggested that one or a small number of mutations may be involved in the occurrence of the DM phenotype. In fact, the determination of the molecular basis of DM showed that the expansion of an unstable CTG repeat sequence is responsible for nearly all cases of DM. Different probes for this genomic region (cDNA25, pBBO.7, pGB2.6, pGP1.5) all identify the same 1 kb insertion/deletion polymorphism (D19S95) detected 5 kb upstream from the unstable CTG region with various restriction enzymes (EcoRI, BamHI, EcoRV, KpnI, and HindIII). Another 1 kb insertion is present in all DM patients with the CTG expansion genotyped to date.2,7,17 Another distinct BamHI polymorphism detected by a more distant probe (p37.1) showed strong disequilibrium with DM. Altogether these data provide further evidence for a founder chromosome effect in DM.

DM and fragile X syndrome have features in common in that they both result from an amplification of a trinucleotide repeat that expands to several thousand copies in the most extreme cases and provide evidence for a founder effect. Therefore, to check whether or not a unique founder chromosome was also involved in French DM, we analysed a large sample of French DM families. We have identified the mutation in 87 French DM families and haplotyped the affected subjects using probe cDNA25 and two markers flanking the DM locus, D19S63 and D19S51.

Material and methods

FAMILIES

A total of 772 subjects from 87 families with myotonic dystrophy was studied. All were of white French origin. The presence or absence of DM was evaluated clinically in all cases, according to the diagnostic criteria described by Griggs et al, by electromyography, slit lamp examination, or muscle biopsy when necessary for questionable cases. Only persons whose status could be unequivocally determined were used in this study.
SOUTHERN BLOT EXPERIMENTS
DNA was prepared from whole blood as previously reported.9 Restriction enzyme digests were performed according to the manufacturers' instructions and fragments were resolved by 1% agarose gel electrophoresis. DNA was transferred onto Hybond-N membranes (Amersham). Probes were labelled with ($\alpha^{32}$P) dCTP (C strategically) and visualised (Lyodean, et al26). Detailed information on the molecular probes and RFLPs tested is given in table 1. The two polymorphic systems can be distinguished using cDNA25 and BamHI, because this enzyme cuts between the site of the 1 kb insertion and the CTG repeat (figure). The largest amplifications of the CTG repeat are more readily visualised in EcoRI digests.

STATISTICAL METHODS
Allele frequencies were estimated by allele counting. All haplotypes were determined by family studies.

Results and discussion
Allele frequencies and PIC values of the three DNA polymorphisms studied in our French population (cDNA25/BamHI, D19S63/ PvuII, D19S51/BglI) are given in table 2. Because there was no statistically significant difference for the non-DM chromosome allele frequencies between DM patients and non-DM subjects, we pooled data from all non-DM chromosomes to calculate allele frequencies at each locus. No significant difference in allele frequencies at these three loci was observed between our French non-DM population and the other white populations studied.1,16

We observed significant allele frequency differences at the D19S63/PvuII locus between DM and non-DM chromosomes (p<0.001). The frequency of allele 3 (D19S63/PvuII) was significantly higher in the DM chromosomes in the French population as described previously in other populations, namely British,15 Spanish,16 Japanese, and Israeli.28 The allele most frequently associated with DM in other populations was allele 3 whereas in the French population allele 1 remained the one most frequently associated with DM. The frequency of allele 3 in DM chromosomes, 0.37 in the French, was lower than in British (0.59), Spanish (0.48), Israeli (0.62), and Japanese (0.50) DM populations.15,16,28 We did not detect any significant difference at the D19S51/BglI locus in our French population. In contrast, Tsilifidis et al20 found significant linkage disequilibrium between the rare D19S51/BglI allele and DM in an extensive linkage analysis of a French-Canadian population. This observation can be explained by a founder effect in the French-Canadian population. Our DM families are derived from a much broader European population of mixed origin and therefore are not subject to such an effect.

Most striking, as shown in table 2, is the global finding that, in DM patients, the mutation is associated with the same cDNA25/BamHI allele (1), while in the normal population alleles 1 and 2 were equally distributed (0.50 each) (p<0.001). As shown in the figure, probing BamHI digests with cDNA25 detects a 1.4 kb fragment that contains the CTG repeat and two alleles of 4.6 and 5.6 kb. The latter contains the 1 kb insertion also detected in an EcoRI digest (10 kb allele). The BamHI 1.4 kb fragment varies in size in DM patients. Small increases in the 1.4 kb BamHI fragment could be observed in about one-third of patients whose alterations could not be resolved in EcoRI digests (Lavedan, et al, unpublished data). BamHI digests are thus complementary to EcoRI digests, and allow one to resolve the two different polymorphisms, namely the 1 kb insertion and the CTG repeat.26 Using BamHI in addition to EcoRI for genetic phasing in French DM families, we have shown unambiguously that the 10 kb EcoRI allele, corresponding to the 5.6 kb BamHI allele, is the expanding allele in all 72 families analysed with cDNA25. This same absolute linkage disequilibrium between the DM mutation and the insertion/deletion polymorphism was observed in previous studies in both white and Japanese patients.27,17

The polymorphisms at the three different loci were used to construct compound haplotypes for DM and non-DM chromosomes in our 72 French DM families (table 3). Some haplotypes (3, 4, 7, 8, 11, 12) were never present in DM patients. This reflects the absolute linkage disequilibrium between the DM mutation and the insertion/deletion polymorphism. Haplotype 2 is seen in 32% of the DM

Restriction map of the genomic segment containing polymorphisms detected at the DM locus. Triangles are for the normal insertion polymorphism (1 kb) and the variable length polymorphism (CTG). The orientation of the map is given from centromere (cen) to telomere (tel). B and E show the restriction site positions for BamHI and EcoRI respectively. The lengths indicated are for chromosomes lacking the insertions. Also indicated are the locations for D19S63 and D19S51.

Table 1 List of DNA polymorphisms studied.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Probe</th>
<th>Enzyme</th>
<th>Allelic fragments (kb)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>D19S63</td>
<td>pD10</td>
<td>PvuII</td>
<td>7.1/6.8/6.5</td>
<td>Brook et al9</td>
</tr>
<tr>
<td>DM</td>
<td>cDNA25</td>
<td>BamHI</td>
<td>5.8/4.6</td>
<td>Shelbourne et al9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EcoRI</td>
<td>9.6/8.6</td>
<td>Buxton et al9</td>
</tr>
<tr>
<td>D19S51</td>
<td>p134C</td>
<td>BglI</td>
<td>5.0/4.4</td>
<td>Johnson et al9</td>
</tr>
</tbody>
</table>

*Note: Information on the molecular probes and RFLPs tested is given in table 1. The two polymorphic systems can be distinguished using cDNA25 and BamHI, because this enzyme cuts between the site of the 1 kb insertion and the CTG repeat (figure). The largest amplifications of the CTG repeat are more readily visualised in EcoRI digests.*
Table 2 DM and non-DM chromosome allele frequencies in different populations.

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Enzyme</th>
<th>Allele</th>
<th>DM patients</th>
<th>DM and non-DM subjects</th>
<th>Other populations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>DM chromosomes</td>
<td>Non-DM chromosomes</td>
<td>Both chromosomes</td>
</tr>
<tr>
<td>D19S63</td>
<td>PeuI</td>
<td>1</td>
<td>0.45 (31)</td>
<td>0.60 (41)</td>
<td>0.58 (98)</td>
</tr>
<tr>
<td>2</td>
<td>0.18 (12)</td>
<td>0.28 (19)</td>
<td>0.29 (49)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.37 (25)</td>
<td>0.12 (8)</td>
<td>0.13 (22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM</td>
<td>BamHI</td>
<td>1</td>
<td>1.00 (72)</td>
<td>0.53 (38)</td>
<td>0.49 (78)</td>
</tr>
<tr>
<td>(cDNA25)</td>
<td></td>
<td>2</td>
<td>0.00 (0)</td>
<td>0.47 (34)</td>
<td>0.51 (82)</td>
</tr>
<tr>
<td>D19S351</td>
<td>BglII</td>
<td>1</td>
<td>0.19 (14)</td>
<td>0.17 (13)</td>
<td>0.18 (32)</td>
</tr>
<tr>
<td>2</td>
<td>0.81 (61)</td>
<td>0.83 (62)</td>
<td>0.82 (142)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Numbers in brackets refer to the number of chromosomes analysed.

1 British and French Canadians.13
2 Spanish.18
3 French Canadians.20
4 Japanese.21
5 Finns.16
6 chromosomes, as a consequence of the presence of the 1 kb insertion in all DM chromosomes and its more frequent association with allele 1 (D19S63/PeuI).

These data further support the hypothesis of a unique founder chromosome. However, the hypothesis of recurrence of the DM mutation on a chromosome carrying the 1 kb insertion cannot be rejected because of the high frequency of this allele in the general population (~0.50). More informative markers are thus needed to determine whether a few founder chromosomes can be characterised in different ethnic groups as in the fragile X syndrome.22

Morton and Macpherson25 developed a three step model involving four alleles for the fragile X syndrome, to account for the complexity of the molecular observations: inserts are either normal (N), rather small and stable (S), larger and unstable (U), or large (L). This model implies a conversion from N to S, from S to Z, and from Z to L, with different conversion probabilities in the two sexes. This multistep model could apply to myotonic dystrophy: L would correspond to congenital DM transmitted only by some women. This is supported by the difference of the intergenerational Ctg amplification observed between male and female transmission, as recently described.10113 However the mechanisms triggering these conversions are still unknown. In this context, the presence of the 1 kb insertion in all DM chromosomes analysed so far is intriguing. This intragenic insertion, which is frequent in the general population, may be an unstable element. As such it may be one of the necessary steps for the progression of the DM mutation in the multistep process. It has to be assumed that the Ctg repeat repeat when associated with the 1 kb insertion (ancestral chromosome) must remain dormant for several generations. Accordingly, Höwel et al23 have reported several pedigrees with evidence of genealogical connections between families indicating the previous presence of the mutation which had been transmitted through several unaffected subjects. Recently, Imbert et al23 studied in a normal population the association of alleles of the Ctg sequence with alleles of the insertion/deletion polymorphism and of a microsatellite marker located 90 kb from the DM mutation. They found the same linkage disequilibrium between both the frequent (Ctg)10 allele and the large (Ctg)10 allele and the insertion/deletion polymorphism as that observed on DM chromosomes. They suggested that the initial predisposing event consisted of a transition from a (Ctg)10 allele to an allele with 19 to 30 repeats and that alleles (Ctg)9 to (Ctg)10 might act as a reservoir for recurrent DM mutations. The Ctg expansion beyond a certain threshold, approximately 50 repeats, would confer a rapid and irreversible instability, leading to increased severity or earlier onset of the disease (anticipation), or both. Alternatively, the larger insertion allele or its environment may confer a predisposition to the expansion of the Ctg repeat. Therefore, it will be important to understand the potential implication of the insertion fragment and its flanking sequences in the Ctg repeat sequence instability.

This work was supported by grants from the Association Française contre les Myopathies, the Caisse Nationale d’Assurance Maladie des Travailleurs Salariés, and INSERM. Work at Charing Cross was supported by the Muscular Dystrophy Group, the Muscular Dystrophy Association/Piton Foundation through its task force in genetics and by the Central Research Fund of the University of London. We are grateful to Dr D Shaw who kindly provided probe P010. We are also grateful to Drs D Amman, S Aymer, J Battin, D Boggio, J Boué, M L Briard, A Carpenter, M L Chauvet, H Cordier, J Cousin, A David, B Ferrand, M Farcou, E Floet, S Garel, S Gilgengerantz, F Giraud, M Groessen, P Guibaud, J-P Harpey, A Huriez, A Jeanneret, J Kaplan, L Largen-Pier, M Legrand, B Le Marcé, M Le Merrer, N Lepeiré, S Manouesse, M Mathieu, J F Mattei, C Morainé, M Morini, J-C Netter, A Nivelen, A Noir, E Ollagnon, P Parent, M O Peter, N Philip,
C Piussan, F Samson, R Saura, F Serville, M Tournarie, A Touitou, P Tron, J Vigneron, and A Villa for referring the families. This work was done with the technical support of Genethon.