INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Université d'Ottawa • University of Ottawa

Analysis of the Unstable Mutation Responsible for Myotonic Dystrophy

A Thesis Submitted to the School of Graduate Studies

University of Ottawa

In Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy

Department of Microbiology and Immunology

Faculty of Medicine

By

Juana M. Barceló

© Juana M. Barceló, Ottawa, Canada, 1997

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced with the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-20990-3

Canadä

ABSTRACT

Myotonic dystrophy (DM) is an autosomal dominant genetic disease which affects approximately 1 in 8000 individuals globally. This is a multisystemic disorder which primarily targets muscle tissues. The genetic defect underlying DM is a highly unstable trinucleotide CTG repeat sequence located in the 3' untranslated region of a gene encoding a protein with serine/threonine protein kinase activity. The number of CTG repeats in non-DM individuals ranges from 5 to 35, whereas in DM individuals it can range from 50 to over 2000. The mechanism of disease and the role of the kinase are currently unknown.

The main characteristics of the mutation are its expanded length and high instability. The instability usually leads to an increase in the number of CTG repeats as the mutant allele is transmitted from one generation to the next. This is seen concurrently with an increase in the severity of the clinical phenotypes through successive generations.

The mechanism(s) underlying the unstable behaviour of this mutation have been unknown. I studied several possible elements that might affect the observed instability.

These studies were performed on a large number of intergenerational transmissions of the mutant allele in DM patients, as well as on different tissues from the same patients. *In vivo* studies revealed that both the gender of the transmitting parent and the size of

the mutation had a significant effect on its intergenerational dynamics. Instability was also seen in cells that underwent mitosis and meiosis, as well as in patients' cells with very low mitotic activity. In addition, *in vitro* studies showed instability of this sequence in cells from DM patients which were grown in tissue culture. This instability was seen to not necessarily be associated with the cell-cycle-coupled DNA replication in those cells. Since my previous analyses had ruled out recombination between homologous chromosomes as a major element involved in the instability I focused my studies on tests for the possibility that DNA repair was associated with the instability of this mutation. These analyses revealed that there were breaks or gaps occurring specifically within the mutation. Since breaks and gaps are DNA repair intermediates this suggested that the high instability of this mutation in DM patients was due to repair attempts on the structure adopted by the long arrays of CTG repeats.

The results of the previous studies on intergenerational transmissions of the mutation, as well as *in vivo* and *in vitro* studies of the unstable properties of this sequence are also compatible with a DNA repair model of repeat instability.

These data point the way to a more focused approach to the identification of specific mechanistic pathways underlying the expansion of this highly unstable CTG repeat in DM. The results of this research may also help in defining the mechanism(s) of expansion of the mutations of an increasing number of human genetic diseases found to be caused by highly unstable trinucleotide repeats.

ACKNOWLEDGEMENTS

My research would have been impossible without the help of many people. I thank the Department of Microbiology and Immunology, for welcoming me into their graduate program, and all the professors at the University of Ottawa who shared my enthusiasm for this research. Drs. Ken Dimock and Lionel Filion eased my way through both research and administrative hurdles as graduate studies co-ordinators. Drs. Earl Brown, Chaim Birnboim, and Alex MacKenzie, were instrumental as thesis research advisors. Dr. Birnboim's guidance has been invaluable. Dr Mackenzie offered generously his time and feedback through my research pursuits. I thank all my colleagues for their friendship and collaboration. Both the Molecular Genetics and the Cytogenetics Diagnostic Laboratories at the Children's Hospital of Eastern Ontario were generous providers of samples and protocols. My relatives and friends provided unconditional support for my work. I will always be grateful to my research supervisor, Dr Robert Korneluk, who made it all possible. His guidance, support, generosity, humour, and patience, made the experience not only invaluable but enjoyable. I also want to thank the Medical Research Council of Canada and the School of Graduate Studies and Research at the University of Ottawa for financial support.

TABLE OF CONTENTS

Abstract ii

Acknowledgments iv

Table of Contents v

List of Tables x

List of Figures. xi

List of Abbreviations xiii

CHAPTER I

INTRODUCTION: THE BASIS OF MYOTONIC DYSTROPHY AND INSTABILITY OF ITS MUTATION

Clinical presentation and epidemiology . . . 1

Cloning the DM mutation 3

Characteristics of the DMK gene 4

Characteristics of the DMK gene product . . . 5

What causes DM? 6

Known highly unstable trinucleotide repeat sequences in the human genome 9

Simple repetitive sequences in the human genome . . 12

Towards an understanding of the mechanistic basis for the instability of the CTG repeat in DM . . . 15

CHAPTER II

ANALYSIS OF THE DM KINASE CTG REPEAT SEQUENCE

1. Introduction 20

2. Methodology 21

3. Results 27

Size distribution and transmission of normal alleles 27

Size distribution and transmission of mutant alleles 31

Study of possible interactions between normal and mutant alleles 33

4. Discussion 35

CHAPTER III

ANALYSIS OF THE EFFECTS OF THE TRANSMISSION OF THE MUTATION AND ITS INTERGENERATIONAL STABILITY

1. Introduction 39

2. Methodology 41

3. Results 43

Intergenerational amplification of the DM mutation 43

Size of the mutation in offspring as a reflection of the size of the mutation in parents 43

Mutation frequency and protomutation stability 45

Transmission of the DM protomutation 48

The stability of the DM protomutation is not due to interruptions of perfect CTG repeat arrays 50

4. Discussion 52

CHAPTER IV

PARENT OF ORIGIN DIFFERENCES IN EXPANSION OF THE DM MUTATION AND THE CASE OF CONGENITAL DM (CDM)

1. Introduction 59

2. Methodology 60

3. Results 62

General parental-gender effects on the transmission of the mutation 62

Maternal effects in the transmission of the mutation: correlation between CTG repeat length in offspring and frequency of CDM 66

Additive influence of maternal and offspring DMK gene CTG repeat lengths in the genesis of CDM 67

4. Discussion 68

Inheritance of CDM from mothers 71

Parent of origin differences in the outcome of transmission of the DM mutation and implications for the mechanism of expansion of the CTG repeat 73

CHAPTER V

CHARACTERISTICS AND BEHAVIOUR OF THE MUTATION IN TISSUES FROM DM PATIENTS

1. Introduction 75

2. Methodology 76

3. Results 77

Tissue distribution of mutant allele sizes 77

Mitosis or time-dependent variation in the size of the mutation 79

Possible effect of meiosis on mutation size changes: size distribution of the mutation in sperm cells 79

Possible effect of mitosis on mutation changes: tissue culture study 81

4. Discussion 85

CHAPTER VI

THE CELL-CYCLE-LINKED DNA-REPLICATION-INDEPENDENT MECHANISM OF EXPANSION OF THE DM MUTATION: A STUDY OF THE POSSIBILITY OF DNA REPAIR AS THE BASIS FOR INSTABILITY

1. Introduction 91

2. Methodology 93

3. Results 96

Cell cycle analysis 96

A cell-cycle-linked DNA-replication-independent mechanism of expansion of the DM mutation 98

Detection of DNA repair intermediates within the mutation 102

Differentiation between single-stranded breaks (SSBs) and double-stranded breaks (DSBs) 104

4. Discussion 106

Finding DNA breaks 106

Recognition and processing of DNA damage 110

A model to explain the instability of the DM mutation 112

CONCLUSIONS 116

REFERENCES 118

APPENDIX 137

LIST OF TABLES

CHAPTER I

1-1 Known highly unstable triplet repeats in humans . . . 10

CHAPTER III

3-1 Summary of amplification results over a range of parental allele sizes 51

LIST OF FIGURES

CHAPTER II

2-1 Location and size distribution of the CTG repeat sequence in the normal allele through polyacrylamide gel electrophoresis . . . 28

2-2 Analysis of the distribution of normal size DM alleles . . 29

2-3 Appearance of normal and mutant alleles following agarose gel electrophoresis of PCR products and Southern blotting . . 30

2-4 Map of genomic region encompassing the CTG repeat in DM . 32

2-5 Examples of autoradiographs of two possible Southern blotting-probing systems used in order to analyse the level of expansion of the CTG repeat sequence as well as its heterogeneity 34

CHAPTER III

3-1 Transmission of the mutation in a DM kindred . . . 44

3-2 Appearance of the CTG repeat following PCR amplification of the mutation in DM individuals with a low number of repeats in the mutant alleles . . 46

3-3 Example of a kindred exhibiting intergenerational stability of the protomutation 47

3-4 Number of repeats inherited from DM parents with mutant allele sizes ranging from 50 to 100 repeats . . 49

3-5 Genealogical reconstruction of DM in the Saguenay-Lac-Saint-Jean region 54

CHAPTER IV

4-1 Maternal and paternal inheritance of the mutation . . 63

4-2 Mean number of repeats present in DM mothers and fathers and increase over this number acquired by their DM offspring 64

4-3 Relationship between maternal and offspring number of repeats and CDM 69

CHAPTER V

5-1 Southern blot analysis of the status of the mutation in tissues from a fetus, a newborn infant, and an adult . . 78

5-2 Expansion of the DM mutation in patient's leukocytes over a time period 80

5-3 Comparative PCR analysis of mutant allele sizes in an individual's blood leukocytes, sperm cells, and blood leukocytes in his offspring 82

5-4 Comparative genomic Southern blot analysis of mutant allele sizes in an individual's blood leukocytes, sperm cells, and blood leukocytes in his offspring . . 83

5-5 Study of the behaviour of the mutant allele in tissue culture 84

CHAPTER VI

6-1 Cell cycle profile obtained through FACS analysis of non-syncronous cycling and confluent (synchronous at G0) fibroblasts 97

6-2 Analysis of expansion of the mutation through active cell cycling and through contact inhibition . . . 99

6-3 Analysis of the extent of expansion acquired by cycling and contact-inhibited cells 101

6-4 Detection of breaks in the DNA cellstreated with DNA repair inhibitors 103

6-5 Detection of double-stranded breaks (DSBs) in DNA of repair inhibitor treated cells 105

LIST OF ABBREVIATIONS

Ara A	adenine arabinoside
Ara C	cytosine arabinoside
bp	base pair
°C	degree Celsius
CDM	congenital DM
cDNA	complementary DNA
cm	centimetre
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
DM	dystrophia myotonica/myotonic dystrophy
DMK	DM kinase
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DSB	double-stranded break
dsDNA	double-stranded DNA
dTTP	deoxythymidine triphosphate
EDTA	ethylenediaminetetraacetic acid
FACS	fluorescent activated cell sorting

gp32	gene 32 protein
h	hour
kb	kilobase
kbp	kilobasepair
kDa	kilodalton
l	litre
LCR	locus control region
mA	milliamperes
min	minute
ml	millilitre
mM	millimolar
mmol	millimole
mRNA	messenger ribonucleic acid
μg	microgram
μl	microlitre
μM	micromolar
NEB	New England Biolabs
ng	nanogram
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEG	polyethylene glycol
pg	picogram

RNA	ribonucleic acid
RT	room temperature
SDS	sodium dodecyl sulfate
sec	second
SSB	single-stranded break
SSC	standard saline citrate
ssDNA	single-stranded DNA
TBE	Tris-borate EDTA
TE	Tris-HCl EDTA
U	unit
uv	ultraviolet
UTR	untranslated region
V	volts
VNTR	variable number of tandem repeats
YAC	yeast artificial chromosome

CHAPTER I

INTRODUCTION: THE BASIS OF MYOTONIC DYSTROPHY AND INSTABILITY OF ITS MUTATION

Clinical presentation and epidemiology

Myotonic dystrophy (DM), also known as Dystrophia Myotonica and Steinert disease, is the most common inherited adult form of neuromuscular disorder, affecting approximately 1 in 8000 individuals globally (Harper, 1989). Canadians are known to have one of the highest incidences of this disease, with frequencies as high as 1 in 529 in the Saguenay-Lac-Saint-Jean region of Quebec (Mathieu et al., 1990). DM is inherited in an autosomal dominant fashion, whereby only one of the two alleles needs to be mutated for disease manifestation to occur. This disease is multisystemic and its primary characteristics are myotonia and muscle weakness, although neurological, cardiac, respiratory, endocrine, and ophthalmological manifestations are also common. Both the phenotype of individuals with DM and the expressivity and penetrance of the disease are quite variable (Harper, 1989). There are two distinct clinical manifestations of DM: an adult onset form and a congenital form (CDM), which is present at birth.

The first symptoms of the adult form of DM may appear in adolescence or later in life. Myotonia (delayed muscle-relaxation capacity) appears as the prominent symptom and

is often followed by premature balding and testicular tubular atrophy (in males), as well as cataracts, retinal degeneration, cardiac conduction defects, diabetes and calcifying epithelia (Harper, 1989).

CDM is a more severe form of DM, differing significantly from the adult form (Vanier, 1960). CDM is obvious at birth, and is often fatal at this stage. Clinical features include extreme hypotonia, muscular atrophy, facial paralysis, respiratory distress, and feeding difficulties. Mental retardation is a common characteristic in infants able to survive the neonatal period. This form of the disease was observed to be exclusively transmitted by mothers with DM although there have been reports suggesting rare cases of CDM from paternal transmission (Nakagawa et al., 1994; Ohya et al., 1994; Koga et al., 1994). Two prominent features of adult onset DM, myotonia and cataracts, are characteristically absent in CDM (Harper, 1989).

In addition to the wide range of clinical symptoms, the age of onset of this disease can also be quite variable. Whereas symptoms of CDM appear at birth, and often lead to death at this stage, the adult form of DM can appear at any stage later in life and only be manifested by cataracts and/or mild myotonia. The number and severity of disease symptoms is quite variable in the different individuals, even among those of the same pedigree.

Genetic anticipation is another characteristic of this disease, and is defined as an

increase in disease severity following transmission of the mutation from one generation to the next. In DM pedigrees this is manifested as a decrease in the age of onset together with an increase of disease severity and symptoms following intergenerational transmission of the mutation (Howeler et al., 1989).

Cloning the DM mutation

No clues to the biochemical basis for DM were available prior to the identification of the mutation. Positional cloning was performed in order to study the putative area of the genome in search of possible candidate genes (reviewed in Shutler, 1993, PhD thesis). Initial linkage studies had mapped the genetic defect to the long arm of chromosome 19 (Hulsebos et al., 1985; Myklebost and Rogne, 1988; Smit et al., 1988). Genetic and physical mapping identified a region bracketed by two genetic markers, ERCC1 (Smeets et al., 1989) and X75b (Jansen et al., 1992a) from which a contig (of 700 kbp) was constructed. This contig consisted of overlapping cosmids and YACs (Shutler et al., 1992; Aslanidis et al., 1992; Jansen et al., 1992a). Systematic screening of likely gene candidates within those sequences was performed. The genes were sequenced in search of base substitutions, insertions and/or deletions specific to DM patients. The search culminated with the discovery of an unstable CTG repeat sequence associated with DM. An amplified CTG repeat segment was found to be specific to individuals with the disease. Initially, only approximately 70% seemed to show this mutation (using genomic Southern blot analysis) in one of their chromosomes (Harley

et al., 1992b; Buxton et al., 1992; Aslanidis et al., 1992). As a result, using primers flanking the mutation, PCR was performed with the rest of the obligate carriers. A mutant allele, albeit of minimal expansion, was detected in these individuals. Sequence analysis of the region flanking the mutation revealed the location of these CTG repeats in the 3' untranslated region (UTR) of a gene. This gene was found to have significant homology to a family of genes encoding serine/threonine protein kinases (Brook et al., 1992; Fu et al., 1992; Mahadevan et al., 1992) and is known as the DM kinase or DMK.

Characteristics of the DMK gene

The structure of the DMK gene consists of approximately 14 kbp of unique sequence (GeneBank accession number L08835) containing 15 exons (ranging in size from 47 to 936 bp), two Alu repeats preceding exon 9, and three Alu repeats following this exon (Mahadevan et al., 1993a). These Alu repeats have been referred to as the 1 kb insertion/deletion polymorphism within this gene and are located 5.3 kb upstream of the CTG repeat. The introns vary in size from 2331 bp (first intron) to 76 bp (Mahadevan et al., 1993b; Figure 2-4). The human DMK sequence has high homology (>90%) to the murine gene (Jansen et al., 1992b; Mahadevan et al., 1993b). Both homologues seem to present with a variety of alternatively spliced mRNAs (Jansen et al., 1992b) and the predicted full length mRNA of the human gene is approximately 3.4 kb. The putative translational start codon was identified based upon the maintenance of the open

reading frame for cDNAs previously identified in both species. No clear TATA or CAAT box motifs were found to be present in either species but computer analysis revealed the presence of putative binding sites for transcription factors in the putative promoter and first intron regions (Mahadevan et al., 1993b).

Characteristics of the DMK gene product

The predicted full-length human DMK gene encodes 629 amino acids (Jansen et al., 1992b; Mahadevan et al., 1993b). Potential structural coding domains of this gene include: a leucine zipper domain in exon 1, a cAMP-dependent serine/threonine protein kinase activity domain in exons 2 to 8, and a coiled-coil region resembling that found in the myosin heavy chain in exons 9 to 12. Exons 13 and 14 do not show any homology to known structures but exon 15 codes for a hydrophobic region in the protein that could act as a transmembrane domain. The murine homologue of the human DMK gene shares 86% identity and 90% similarity at the amino acid level. The exact function of the DMK protein is unknown. Human DMK, as identified by a polyclonal antibody against a DMK fusion protein, appears as two distinct protein species of approximately 71-74 and 80-82 kDa (Whiting et al., 1995a; Maeda et al., 1995). This protein was localized at the postsynaptic neuromuscular junction in skeletal muscle, to intercalated discs in cardiac tissue, and to ventricular ependymal cells and choroid plexus in the brain (Whiting et al., 1995a; van der Ven et al., 1993). The mutation in the 3' UTR is not expected to directly affect the amino acid sequence but alterations

in mutant DMK could arise from different spliced versions of this protein. Although the function of this protein is currently unknown, recent studies suggest that the DMK is not membrane-integrated although it may have a molecular organization which favours peripheral association with membranes (Waring et al., 1996). This preferential association is predicted to be mediated through the C terminus, and the protein appears to participate in the formation of high molecular weight disulphide-linked complexes (Waring et al., 1996).

What causes DM?

There is consensus on the causal relationship between the existence of an amplified CTG repeat in the 3' UTR of the DMK gene and the disease. However, there is no such consensus when it comes to explaining how the disease is induced by the mutation. Studies of the expression of DMK in tissues from patients are controversial and differ in their results. Most studies report a DMK haploinsufficiency mediated model of disease induction. This has been proposed to be caused by decreased DMK gene expression levels, either through decreased expression of steady state DMK transcript levels (Fu et al., 1993; Novelli et al., 1993; Carango et al., 1993; Tachi, 1996), through lower levels of mutant gene transcripts (Carango et al., 1993; Hoffman-Radvanyi et al., 1993) or as decreased levels of DMK protein in patients (Fu et al., 1993; Koga et al., 1994; Maeda et al., 1995). In contrast, Sabourin et al., (1993), reported increased DMK gene mRNA levels in tissues originating from an infant suffering from CDM. More

recent studies, in which allele specific quantitative reverse transcription PCR was performed, seemed to suggest no change (Krahe et al., 1995a) or a small decrease (Wang et al., 1995) in total DMK transcripts. The use of various antibodies against the DMK protein suggest either a reduction (Fu et al., 1993; Koga et al., 1994; Maeda et al., 1995), an increase (Dunne et al., 1996), or practically no change (Sabourin, 1996, PhD thesis) in the levels of the DMK in tissues in affected patients.

Models of human disease in mice, even with the undisputed gene alterations known to cause disease in humans, are not always successful. Lack of the expected phenotype does not necessarily show that the putative gene defect is not the cause of disease. Nevertheless, reports of DMK knockout (Jansen et al., 1996; Reddy et al., 1996) and transgenic mice (Jansen, et al., 1996; Narang et al., submitted manuscript) seem to suggest that this protein can play a role in maintaining skeletal muscle function. Moreover, loss of the DMK gene is associated with very little, if any, phenotypic changes. Indeed, none of these mouse models recreate the multisystemic DM phenotype. As a result, investigators are questioning the hypothesis that alterations of the DMK alone (whether it is through under- or over-expression or by affecting alternative splicing) could be responsible for DM (Wieringa, 1994; Harris et at., 1996; Hamshere and Brook, 1996). Indeed, the DM kinase may not be the sole gene that is affected in DM, that is, the only one in which expression is affected by the presence of the expanded CTG repeats. The transcription of other genes in cis to the mutation may be affected in DM patients.

A model of disease phenotype induction, mediated by the DMK mRNA, has been proposed. In this model the disease would be mediated through altered nuclear transport, or impaired mRNA processing in the nucleus (Sabourin, 1996, PhD thesis). Accumulation of nuclear foci, apparently corresponding to the mutant allele mRNA transcript, was detected in fibroblasts from DM patients (Taneja et al., 1995). It has been proposed that the 3' UTR of the mutant transcript could become a sink for nuclear proteins involved in a variety of processes and in this way induce the disease. While these and other hypotheses based on the toxic accumulation of DMK gene mutant mRNA transcripts should be tested, the emphasis of some researchers has focused on the characterization of genes in the vicinity of DMK and the study of their possibly altered expression patterns. The DMK gene is also part of a large CpG island associated with a novel downstream gene encoding a putative homeodomain protein (DMAHP, Boucher et al., 1995). Upstream of the DMK gene, at a distance of only 0.5 kb there is another gene, called gene 59 (DMR-N9 in the mouse), (Jansen et al., 1995), whose function is still unknown but which was found to be highly expressed in testes and brain.

There have been multiple reports on the formation of alternative DNA structures *in vitro* by the disease-linked trinucleotide repeats (Fry et al., 1994; Gacy et al., 1995; Mitas et al., 1995; Pearson and Sinden, 1996; Mariappan et al., 1996). Loss of *DNaseI* hypersensitivity *in vivo* at a site located in the proximity of the DM mutation (Otten and Tapscott, 1995) may suggest a disease mechanism mediated by short or long-range

chromatin effects. Histones were reported to bind preferentially to the DM repeats (Wang et al., 1994b; Wang and Griffith, 1995; Godde and Wolffe, 1996). Effects similar to those in heterochromatin-induced position effect variegation in *Drosophila* are suspected to be responsible for some human diseases (Hendrich and Willard, 1995; Milot et al., 1996). It is possible that the exceedingly long tracts of CTG repeats, with their associated *in vivo* structures, would affect the DMK and neighbouring genes at the transcriptional level. Even subtle alterations in the regulation of gene expression in the DM region could possibly induce, through multigene involvement, the multisystemic phenotypic features of this disease. The chromosomal locus of the DMK was seen to be early replicating within the S-phase of the cell cycle (Barceló et al., unpublished results). This observation, and the fact that DM is a dominant disease, are consistent with a model of disease induction in which the mutation disrupts the suspectedly tight transcription control needed at this genomic region.

Known highly unstable trinucleotide repeat sequences in the human genome

To date, fourteen genomic loci in humans have been seen to exhibit expansion of a trinucleotide repeat sequence upon intergenerational transmission. Of those, twelve have been seen to cause disease and five cause a chromosomal perturbation called a fragile site (Table 1-1). At least two of these fragile sites are apparently benign. Instability is manifested as a predilection towards amplification, which translates into

Notes on Table 1-1:

Fragile sites refer to folate-sensitive chromosomal aberrations that appear as poorly condensed regions in metaphase chromosomes (Sutherland and Hecht, 1985).

SBMA = Spinal and bulbar muscular atrophy or Kennedy's disease
DRPLA = Dentatorubral pallidolusyan atrophy
SCA = Spinocerebellar ataxia
HD = Huntington's disease

The first line in the number of repeating units refers to the status in normal individuals, the second line refers to mutation carriers and affected individuals.

References:

Oberlé et al., 1991; Verkerk et al., 1991; La Spada et al., 1991; Mahadevan et al., 1992; Brook et al., 1992; Fu et al., 1992; Brown et al., 1993; Snow et al., 1993; Barceló et al., 1993; Campuzano et al., 1996; Warren, 1996; David et al., 1996; Sanpei et al., 1996; Pulst et al., 1996; Imbert et al., 1996; and various posters and oral presentations at the American Society of Human Genetics Meeting, San Francisco, CA, Oct-Nov, 1996

Table 1-1. Highly unstable triplet repeats in humans

Gene	Location in gene	Repeat unit	No of repeats	Disease	Chromosomal site
FMR1	5′ UTR	CGG	10-55 43->2000	Fragile X syndrome	Xq27.3 (Fragile site Xq27.3)
CBL2	5′ UTR	CGG	8-32 75->1000	Jacobsen syndrome?	11q23.3 (Fragile site 11B)
FMR2	5′	CGG	6-25 50->2000	FRAXE mild mental retardation	Xq28 (Fragile site XM MR)
X25	Intron	GAA	7-22 200->900	Friedreichs ataxia	9q13
AR	Exon	CAG	12-34 40-62	SBMA	Xq11-12
SCA1	Exon	CAG	6-39 41-81	SCA type 1	6p22-23
SCA2	Exon	CAG	17-29 35-59	SCA type 2	12q23-24
SCA7	Exon	CAG	not reported	SCA type 7	3p12-13
Hunting-tin	Exon	CAG	6-37 35-121	HD	4p16.3
DRPLA	Exon	CAG	7-34 54-70	Haw River syndrome or DRPLA	12p12-13
MJD1/ SCA3	Exon	CAG	13-36 68-79	Machado-Joseph disease	14q24.3-32.1
DMK	3′UTR	CTG	5-37 50->2000	DM	19q13.3
unknown	unknown	CGG	10-38 300-500	unknown	unknown (Fragile site XF)
unknown	unknown	CGG	16-49 1000-2000	unknown	unknown (Fragile site 16A)

intergenerational expansions with siblings showing mutations of different size among them and distinct from those of the transmitting parents. Genetic anticipation is a hallmark in these diseases, and is linked to the intergenerational amplification of the trinucleotide repeat. These mutations have been found in different locations in relation to the putative genes involved in disease etiology (Table 1-1). One of the most recently discovered trinucleotide repeat expansions causing human disease is the GAA repeat responsible for Friedreich's ataxia (FA), an autosomal recessive disease. FA is caused by the expansion of this trinucleotide in an intron of the frataxin (or X25) gene (Campuzano et al., 1996). Except for this GAA repeat, all the rest contain either CGG or CAG repeats in either strand. Interestingly, in all of them the expansion occurs in a trinucleotide as opposed to a di- or a tetra- nucleotide, for example. Also, except for FA, these disorders show genetic anticipation, are transmitted as dominant traits (except for SBMA and fragile X syndrome, which are X-liked) and have the repeat mutation in exons of the respective genes. Again, with the exception of the case of FA, all the described unstable repeats have a high GC content in their sequence. Furthermore, all of the trinucleotide repeat expansion mutations are, to a certain extent, involved in the induction of neurological disorders. The biological basis of these coincidences is unknown. The largest expansions of an unstable triplet repeat are found when the mutation is located outside the coding region, as is the case for fragile X and DM. In FA, where the triplet expansion mutation is located in an intron, the number of GAA triplets has been seen to be over 900 (Campuzano et al., 1996). The rest (see Table 1-1) present with moderate expansions of their trinucleotides (usually fewer than 100

repeats). This may reflect a link between transcription and expansion or the fact that transcribed larger expansions constitute embryonic lethals. The trinucleotide repeats in coding regions seem to induce a gain- or change-of-function through the addition of polyglutamine tracts in their respective proteins. In FA and fragile X the mutation seems to lead to a loss of function of the protein.

Simple repetitive sequences in the human genome

The highly unstable sequences described are located at, and are the result of, expansions of already existing shorter arrays of the same trinucleotide repeats at the same sites. These short interspersed simple repetitive DNA sequences are common in the human genome and are polymorphic in length because of variability in the number of tandemly repeated units. These sequences are also found in the genomes of a variety of species of vertebrates, invertebrates, and plants.

One class of repetitive sequences that are highly polymorphic are the VNTR (variable number of tandem repeats) sequences. The number of repeating units in VNTRs ranges from 9 to 40 (Jeffreys et al., 1985; Nakamura et al., 1987). VNTRs are not randomly distributed in the human genome; instead, they are preferentially located in the telomeric regions of chromosomes (Royle et al., 1987; Nakamura et al., 1988).

Microsatellites are another class of tandemly repeated sequences that are more evenly

distributed. They consist of repeated mono-, di-, tri-, tetra-, penta-, or hexanucleotides. The trinucleotide repeats that are associated with human genetic disease have been classified as microsatellites. The most widely studied microsatellite consists of a variable number of the dinucleotide repeat (CA)n where n ranges from 10 to 60. There are estimates of the existence of 50,000-100,000 copies of (CA)n repeats in the human genome (Miesfel et al., 1981; Hamada et al., 1982; Litt and Lutty, 1989), occurring at an average frequency of once every 30 kb (Stallings et al., 1991). Their mean heterozygosity index is 70% (Dib et al., 1996), whereas heterozygosity indices of VNTRs can be as high as 90%.

There is a third class of tandem repetitive sequence which is located adjacent to Alu sequences and is also quite polymorphic. These kinds of sequences have been called "Alu sequence-related polymorphisms" (Zuliani and Hobbs, 1990). Their heterozygosity index is similar to that seen in microsatellites (Weber and May, 1988).

The function, if any, of these repetitive sequences is not well understood. The fact that they are ubiquitous and very polymorphic has converted them into valuable markers in linkage analysis and gene mapping. Their polymorphic information content, manifested as the variable number of repeating units, has led to speculations about possible mechanisms for their generation.

Homologous and unequal crossing-over as well as slipped-strand mispairing within the

repeats during DNA replication have been proposed as likely candidate mechanisms for the generation of their high polymorphism. The estimation of minisatellite mutation rates to new length alleles in human pedigrees has shown that maternal and paternal mutations arise with similar frequency, consistent with length-change being restricted to one stage of gametogenesis, possibly meiosis (Jeffreys et al., 1988). However, analysis of markers flanking new mutant minisatellite alleles failed to show unequal crossing-over between homologous chromosomes for 12 different mutants of two VNTRs (Wolff et al., 1988, 1989). The study of clonal tumour cell populations (Armour et al., 1989) and early mouse embryos (Kelly et al., 1989) indicate that instability of microsatellites is also possible in the absence of meiosis.

Slipped-strand mispairing can occur in the absence of unequal crossing-over of sister chromatids and could potentially occur whenever unpaired loops form, such as during DNA replication or repair. Slippage has been proposed as a possible mechanism of amplification of DNA repetitive sequences (Drake et al., 1983; Streisinger et al., 1966; Wells et al., 1965; Morgan et al., 1974; Efstratiadis et al., 1980). Levinson and Gutman, (1987), proposed that slipped-strand mispairing could be the major force behind the expansion of simple repeat sequences. However, the precise mechanisms by which the microsatellite sequences seen in human genetic diseases (Table 1-1) can be so highly unstable and expanded to such high copy numbers, is still largely unknown and constitutes an area of intensive research.

Towards an understanding of the mechanistic basis for the instability of the CTG repeat in DM

The primary objective of my research was to study possible mechanistic pathways underlying the high instability of the DM CTG repeat. This thesis describes the highlights in this pursuit.

There are various reasons why the definition of the basis of the expansion of the DM mutation is important. First, a high and increasing number of genetic diseases occur due to unstable triple repeats. Hence, finding the mechanism of expansion of the CTG repeat in DM is likely to help understand the same process in the other diseases. Second, the inheritance of highly unstable repetitive DNA, such as that seen in these diseases, is a newly discovered process that challenges the traditional understanding of the inheritance of genetic information. Understanding the basis of the instability of this repeat sequence will help understand aspects related to DNA metabolism and unstable DNA in general. Third, since the expanded CTG repeat in DM underlies the disease, future therapeutic approaches could be derived from understanding the process of expansion. Therapeutic advances towards the control of the disease which are based on targeting the affected gene products (mRNA and/or proteins) would have to await undisputed results on disease mechanisms.

Some of the results of my research into factors linked to the instability of the DM

mutation contributed to the research into other aspects of this disease. For example, studies on how the length of the DM CTG repeat array contributes to the dynamics of the mutation provided molecular data that could help to explain epidemiological aspects of DM. The analyses of parental-gender effects on the intergenerational dynamics of the mutation gave an insight on the possible origin of the distinct phenotypic manifestations of CDM.

It is not clear why the high instability of the described trinucleotide repeat sequences shown in Table 1-1 occurs. Although the composition of these mutations is the same as that of microsatellites, the same instability mechanism(s) may not be shared. Microsatellites in general are not usually found as such long arrays of usually perfect repeating units, and the rate of instability of microsatellites is orders of magnitude lower than that seen, for example, in the DM mutation.

As an initial approach to examining the CTG repeat instability in DM, I studied the events surrounding intergenerational transmission of this mutation. My research started with the characterization of the newly discovered CTG repeat sequence in the DM and non-DM populations (Chapter II; Mahadevan et al., 1992; Neville et al., 1994), followed later by the study of the transmission of the DM mutation in a large population (Chapters III-IV; Barceló et al., 1993, 1994; Tsilfidis et al., 1992; Ashizawa et al., 1994). The size of the mutation refered to in these studies corresponds to that found in peripheral blood leukocytes. This size does not necessarily reflect that of other

tissues (Chapter V). Intergenerational transmission of normal size alleles (<35 CTG repeats) was seen to result in no changes in the number of repeats. Although normal alleles were quite polymorphic in size, their relative stability was also revealed by the fact that they were in linkage disequilibrium with other single base polymorphic markers in close proximity to the mutation (Neville et al., 1994).

While normal alleles were stable, mutant alleles exhibited an amplification following intergenerational transmission which was partly determined by the size of the mutation in the parent. Transmission of the mutation from parents with a low number of repeats (<80) in their blood leukocytes often resulted in little or no changes in the mutation in their offspring (Barceló et al., 1993). Transmission of the mutation from parents with a higher number of repeats was seen to result, in most cases, with a considerable expansion (Chapters III-IV).

Parental-gender related differences in transmission outcome were also observed. As a consequence of one of these studies I obtained evidence of the additive pathology of both maternal and offspring DM CTG repeat array lengths in the development of CDM (Chapter IV; Barceló et al., 1994). Notably, transmission of the mutation by fathers who showed very large CTG repeat expansions in the DNA of their peripheral blood leukocytes was seen to lead, in a high percentage of cases, to reductions in the number of CTG repeats in their offspring (Ashizawa et al., 1994).

The study of tissues from DM patients (Chapter V) revealed age and tissue type linked mutation dynamics. The mutation sequence showed a higher heterogeneity in length in older DM patients. Skeletal muscle tissues in adult patients showed the largest expansions (Anvret et al., 1993; Thornton et al., 1994a; Zatz et al., 1995).

These investigations were followed by tissue culture studies of the mitotic behaviour of the DM mutation. Possible changes in mutation size were monitored in cells originating from DM patients growing in tissue culture. The development of technical improvements allowing for the PCR amplification of DM mutation sequences of any length (Chapter VI; Cheng et al., 1996) led to a more accurate measurement of those changes.

The use of the described cell culture system circumvented the need for cloning the CTG repeat sequence that other systems would have required. (The maximum number of perfect CTG repeats reported to have been integrated in a plasmid is no more than 100). This system was adequate in providing a model for the study of the dynamics of the DM mutation since the mutation sequence remained in its own cellular and genomic context (Chapter V-VI).

Linkage analysis revealed that homologous chromosome recombination did not seem to play a major role in the dynamics of the DM mutation (Neville et al., 1994); therefore, I examined two other major processes in DNA metabolism in relation to the

expansion of this mutation, cell-cycle-coupled DNA replication and DNA repair (Chapter VI). The belief that cell-cycle-coupled DNA replication could underlie the instability of this mutation relies, in part, on models of microsatellite instability that are based on slippage during replication of the repeat units. Indeed, the *in vitro* replication of long stretches of CTG repeats with standard techniques is unusually difficult (Brook et al., 1992; Fu et al., 1992; Mahadevan et al., 1992). Nevertheless, a simple model of slippage during DNA replication could not explain the characteristics of the behaviour of this mutation. After observing expansion of this mutation regardless of cell-cycle-linked DNA replication my research was focused towards the investigation of the possibility of DNA repair as an alternative (Barceló and Korneluk, 1997). This is studied in Chapter VI along with the presence and nature of intermediates of repair processes, such as DNA breaks.

CHAPTER II

ANALYSIS OF THE DM KINASE CTG REPEAT SEQUENCE

1. Introduction

This chapter gives a general overview of my research on the characteristics of the DM microsatellite and basic aspects of its transmission. The methodology section describes the general materials and methods used in DM CTG repeat detection in this and the following chapters.

As noted earlier, the DM mutation consists of an unstable CTG repeat sequence in the 3' UTR of the DMK gene. In DM patients, particularly in those with the most severe clinical symptoms, this repeat sequence is greatly expanded.

Non-DM chromosomes have a range of 5 to 35 CTG repeats at this locus (Brook et al., 1992; Fu et al., 1992; Mahadevan et al., 1992; Davies et al., 1992) that can be detected by standard PCR protocols. These alleles reveal high polymorphism as well as stable intergenerational transmission. The presence of a series of polymorphic markers (base substitutions and an Alu sequence insertion/deletion) within the DMK gene and flanking regions (Neville et al., 1994) allowed further assessment of their long-term relative stability.

Mutant alleles were analysed and their transmission was studied in a large number of DM families. The development of appropriate molecular diagnosis protocols allowed the detection and quantification of the number of CTG repeats. The number of CTG repeats in mutant alleles was seen to range from 50 to over 2000 (Brook et al., 1992; Fu et al., 1992; Mahadevan et al., 1992; Davies et al., 1992). The instability of these alleles was revealed through both their somatic heterogeneity and their intergenerational expansion. Detection of mutant alleles and quantification of the number of CTG repeats was done by PCR and genomic DNA Southern blotting. The same polymorphic markers that were used for the assessment of long-term relative stability of normal alleles allowed the assessment of the possible role of homologous chromosome recombination in the expansion process.

The results of these analyses gave the first indications of factors that were or were not likely to be linked to the high instability of this CTG trinucleotide in DM families.

2. Methodology

Myotonic dystrophy diagnosis

All DM families were ascertained with probands who presented clear clinical symptoms of the disease. Criteria for clinical diagnosis were as described previously (MacKenzie et al., 1989) adhering to published recommendations (Griggs and Wood, 1989).

DM and non-DM population study group

The transmission of the mutation was studied in approximately 500 DM families. These included approximately 2000 individuals. DNA samples were collected from clinically affected individuals, obligate mutation carriers (with or without clinical symptoms of DM), and non-mutation carriers in the same families. Most of the DNA samples corresponded to individuals of French-Canadian origin. The "non-DM mutation" carriers, referred to as normal (or non-DM) individuals in the studies, were often either spouses or siblings of mutation carriers or clinically affected individuals.

Genomic DNA preparation

Genomic DNA used for diagnosis of the disease was normally obtained from peripheral blood leukocytes. DNA was also obtained from a variety of tissues from these patients. DNA extraction was either performed with an automated ABI DNA extractor according to the manufacturer's instructions or manually following the protocol of Birnboim (1992) when from cells in culture. The DNA was reconstituted in TE to a concentration of approximately 500µg/ml and kept at 4°C.

Oligonucleotide synthesis

Oligonucleotides used for PCR amplification, sequencing, and as DNA probes on Southern blots, were synthesized on an automated DNA synthesizer (PCR-Mate, model 391, Applied Biosystems). They were generated according to the manufacturer's instructions using the phosphoramidite method. At the end of the synthesis they were

removed from the solid support and deblocked from the protective group using ammonium hydroxide at 55°C overnight. The oligonucletide DNA was then dried and kept at -20°C until needed, at which point it was dissolved in water to an appropriate concentration for use.

Probes

Preparation of probes used in genomic DNA hybridization. The pGB2.2 (a subclone of pGB2.6, Mahadevan et al., 1992) is a 2.2 kb DNA sequence that maps to the 9-10 kb *EcoRI* fragment that contains the 1 kb Alu repeat insertion/deletion polymorphism (Mahadevan et al., 1993a) and the DM mutation locus (Aslanidis et al., 1992; Mahadevan et al., 1992). The probe 750bpBamHI is a 0.7 kb DNA fragment obtained from a plasmid insert flanked by *BamHI* restriction enzyme sites. This fragment maps to a genomic region within DMK that is flanked by *BglI* and *PstI* restriction enzyme sites and contains the DM mutation locus. The probes were generated through electrophoretic gel fractionation of the corresponding plasmid inserts on a 0.6% low melting temperature agarose (GIBCO) gel, followed by purification of the DNA with glass powder, according to the manufacturer's instructions (BIO 101, Geneclean kit). These probes were radioactively labelled with [α-^{32}P]-dCTP (Amersham). This was done by following the method of Feinberg and Vogelstein (1983), or in a labelling reaction that included the use of *redi*prime® (Amersham), a random primer nucleic acid labelling kit, according to the manufacturer's instructions.

Preparation of probes used for the hybridization of PCR generated DNA. A [γ-^{32}P] (Amersham) labelled $(CAG)_{10}$ oligonucleotide was used as a probe in Southern blots of PCR DNA. This probe was generated by a labelling reaction in which the [γ-^{32}P] of [γ-^{32}P]ATP was linked to a $(CAG)_{10}$ oligonucleotide. The labelling reaction took place at 37°C for 1 h in the presence of the oligonucleotide, T4 polynucleotide kinase, and the appropriate exchange reaction buffer (GIBCO). Ethanol precipitation of the DNA preceded the use of this oligonucleotide as a probe. The alkaline phosphatase conjugated $(CAG)_{10}$ oligonucleotide probe, also used in the hybridization of the same blots, was generated using a commercial kit (E-Link Plus oligonucleotide labelling kit) and following the manufacturer's instructions. Aliquots of this non-radioactive probe were kept at -20°C for up to 1 year prior to use.

Analysis of normal size alleles

Normal alleles were detected through PCR amplification with CTG repeat flanking oligonucleotide primers 409 and 410 (for exact sizing of the repeat sequence in these alleles), or 406 and 409 (Mahadevan et al., 1992; Figure 2-1). The mutation sequence was PCR amplified using genomic DNA as template in a standard reaction mixture. A 50 µl PCR reaction mixture contained 1.5 µl (approximately 750 µg) of genomic DNA, 22 µl of distilled H_2O, 5 µl of 10 x Taq polymerase buffer with 100 mM Tris-HCl pH 8.3, 500 mM KCl, 15 mM $MgCl_2$, 5 µl of each of either one of the oligonucleotide primer pairs 406 and 409 or 409 and 410 (at an initial concentration of 20 µg/µl), 8 µl of a dNTP mix containing the four precursors at a concentration of 1.25 mM each, 2.5

μl of DMSO, and 0.5 μl of Taq polymerase (Cetus, 5 U/μl). When using the primer pair 409 and 410 in the reaction 10 ng of radiolabelled primer 409 were also added. In this case the primer 409 (1μg) was end-labelled with [γ-^{32}P] in a reaction containing [γ-^{32}P]-dATP, T4 polynucleotide kinase, and an appropriate buffer (GIBCO) at 37°C for 30 min. The PCR took place in a Perkin-Elmer thermal cycler 480 using a standard protocol and 25 to 30 cycles of 96°C for 1 min, 60°C for 1 min, and 72°C for 1.5 min, followed by 10 min at 72°C. Products were resolved by denaturing electrophoresis on 8% polyacrylamide urea-containing vertical gels. Gels were exposed to X-AR film (Kodak) at room temperature for 17 hours. Sequencing reactions generated from a control template were used as molecular size standards. The non-radioactive PCR products (reactions containing primers 406 and 409) were submitted to 1-3% agarose (GIBCO) gel electrophoresis for 1-3 h on horizontal gels in 1 x TBE buffer. The products were seen by adding ethidium bromide to the TBE during gel preparation. The 123 bp molecular size marker (GIBCO), loaded in flanking lanes, provided a means of measuring the approximate size and concentration of the products in Polaroid photographs.

Analysis of mutant alleles

Southern blot analysis of PCR DNA. PCR and 1 % agarose gel electrophoresis as described was followed by alkaline (0.4 M NaOH) overnight Southern blot transfer of products onto cationic nylon membranes (Amersham). Membranes were subsequently neutralized in 0.2 M Tris-Cl (pH 7.5) and 2 x SSC, and heat treated to 80°C for 1-2

hours prior to hybridization. Prehybridization took place in a solution containing 5 x SSC, 0.5 % SDS and 5 x Denhardt's solution for 30 min to 1 h at 50°C. Hybridization took place overnight under the same conditions using the identical solution as for prehybridization but also containing 10 ng/ml of [γ-^{32}P] linked to a $(CAG)_{10}$ oligonucleotide as the probe. The membranes were subsequently washed in 6 x SSC and 0.1 % SDS at increasing temperatures up to 45°C, and exposed for 1 h to 20 h at -70°C to X-AR autoradiographic film. Non-radioisotopic mutation detection was performed under the same prehybridizing and hybridizing conditions except that an alkaline phosphatase-conjugated $(CAG)_{10}$ oligonucleotide was used as a probe instead, and hybridization took place for 30 min to 1 h. The membranes were subsequently washed for 5 min in 1 x SSC and 0.5 % SDS at 50°C and for another 5 min in 0.25 x SSC and 0.5 % SDS at 50°C. This was followed by a final wash in 1 x SSC at RT for 5 min. The membranes were then exposed to X-AR autoradiograph film at 37°C for 2 h to overnight.

Southern blot analysis of genomic DNA. Purified genomic DNA (2-5 μg) was digested with *EcoRI*, *BglI*, or *PstI*, and appropriate buffers (GIBCO or NEB) at 37°C for 1 h to 20 h. Reaction products were loaded on 0.6-0.8 % agarose gels in 1 x TBE buffer without ethidium bromide. Electrophoresis took place for 20-22 h at 37 V, 40 mA, and increased to a maximum of 45 V, 50 mA for an additional 5 to 8 h in the case of *EcoRI* digested products. Electrophoresis time was reduced in the case of *BglI* and *PstI* digests. Appropriate molecular size markers (GIBCO) were added to flanking lanes.

Gels were subsequently stained with ethidium bromide and photographed. Both DNA denaturing and Southern blot transfer to a positively charged nylon membrane took place in the presence of 0.4 M NaOH. Membranes were then neutralized and the DNA was then fixed through heat exposure (80°C) for 1-2 h. Membranes were prehybridized at 65°C for 1-3 h in a solution containing 5 x SSC, 0.5% SDS, and 5 x Denhardt's solution. Hybridization took place overnight in the same, but fresh solution and at the same temperature, in the presence of a [α-^{32}P]dCTP-labelled pGB2.2 (or 750bpBamHI) probe sequence. Blots were washed in 0.2 x SSC and 0.1% SDS at 65°C and exposed to X-AR autoradiographic film for 1-4 days. CTG repeat numbers in the mutant alleles were assigned based on the mobility of the mutant compared to the normal allele and a DNA size marker.

3. Results

Size distribution and transmission of normal alleles

Gel electrophoresis (Figures 2-1 and 2-3) and sequence analysis (data not shown) revealed that the different sizes of the normal alleles corresponded to variability in the number of CTG repeats. Polyacrylamide gel electrophoresis (Figure 2-1) allowed the detection of the exact number of repeats in normal alleles. The number of CTG repeats in normal alleles ranged from 5 to 35 in the 240 chromosomes analysed. The CTG repeat numbers 5 and 13 were the most common in the group with a frequency of 35% and 19%, respectively. The size distribution of normal alleles in the population that was

Figure 2-1. Location and size distribution of the CTG repeat sequence in the normal allele through polyacrylamide gel electrophoresis. Analysis of [^{32}P]-labelled PCR products by 8% denaturing polyacrylamide gel electrophoresis showing sequences containing from 5 to 28 CTG repeats. The most common size ranges in the general population (5 repeats, 35%; 13 repeats, 19%) are indicated. The PCR primers (409 and 410) used in the reaction and their locations in relation to the CTG repeats are indicated in the sequence below the photograph. The sequence shown contains 20 CTG repeats and includes the location of some common restriction enzyme sites and two other PCR primers.

Reprinted with permission from *<u>Science</u>* *(Mahadevan et al., Myotonic Dystrophy: An Unstable CTG repeat in the 3' Untranslated Region of the Gene, 255: 1253-1255, 6 March 1992) ®1992 AAAS*

13

5

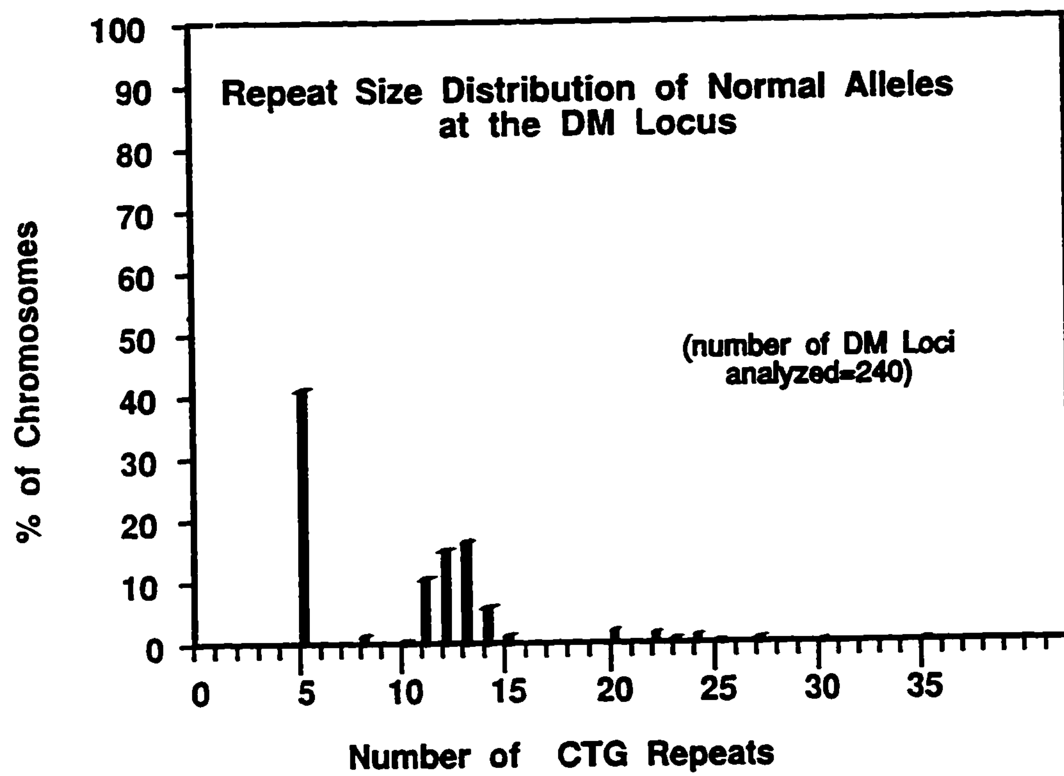
```
    SmaI
    CCCGGGGCACGGCACAGAAGCCGCGCCCACCGCCTGCCAGTTCACAACCGCTCCGAGCGT

 61 GGGTCTCCGCCCAGCTCCAGTCCTGTGATCCGGGCCCGCCCCCTAGCGGCCGGGGAGGGA
                                        primer 409 ►
121 GGGGCCGGGTCCGCGGCCGGCGAACGGGGCTCGAAGGGTCCTTGTAGCCGGGAATGCTGC

181 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGGGG
         ◄ primer 410                                      PstI
241 GATCACAGACCATTTCTTTCTTTCGGCCAGGCTGAGGCCCTGACGTGGATGGGCAAACTG
                                         ◄ primer 406
301 CAGGCCTGGGAAGGCAGCAAGCCGGGCCGTCCGTGTTCCATCCTCCAGCGCACCCCCACC
                          HindIII            ◄ primer 407
361 TATCGTTGGTTCGCAAAGTGCAAAGCTTTCTTGTGCATGACGCCCTGCTCTGGGGAGCGT

421 CTGGCGCGATCTCTGCCTGCTTACTCGGGAAATTTGCTTTTGCCAAACCCGCTTTTTCGG
    BamHI
```

Figure 2-2. Analysis of the distribution of normal size DM alleles. This bar graph reveals the high polymorphism of this microsatellite and its bimodal distribution. Of 240 genes analysed over 40% had 5 and 18% had 13 CTG repeats.

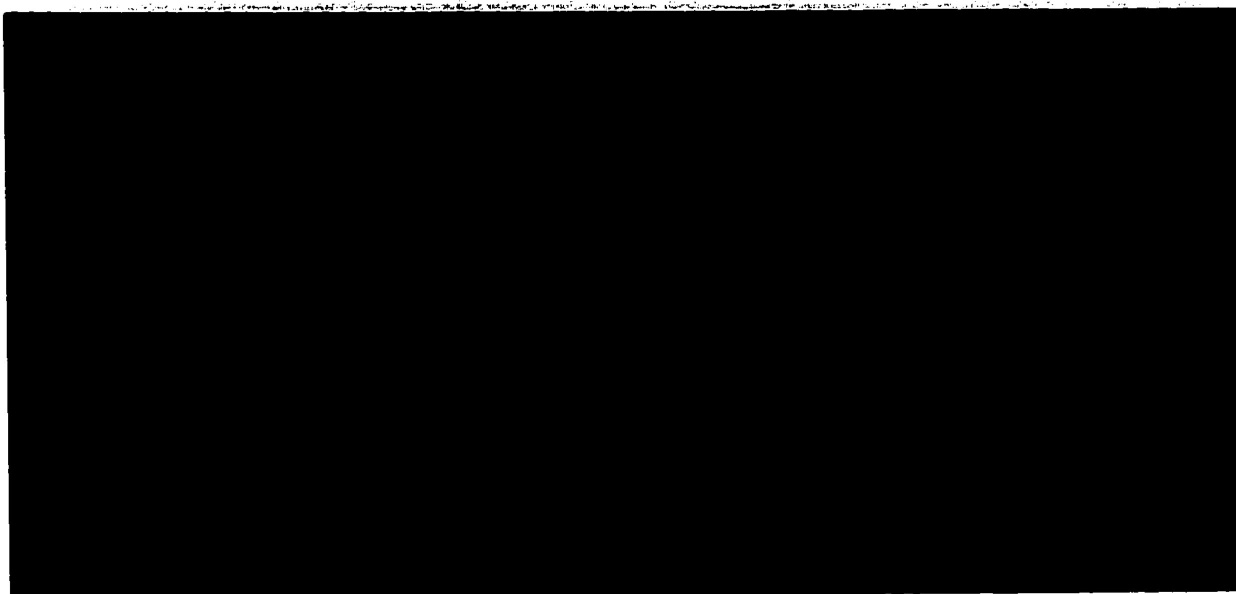


Repeat Size Distribution of Normal Alleles at the DM Locus
(number of DM Loci analyzed=240)
% of Chromosomes
100
90
80
70
60
50
40
30
20
10
0
0
5
10
15
20
25
30
35
Number of CTG Repeats

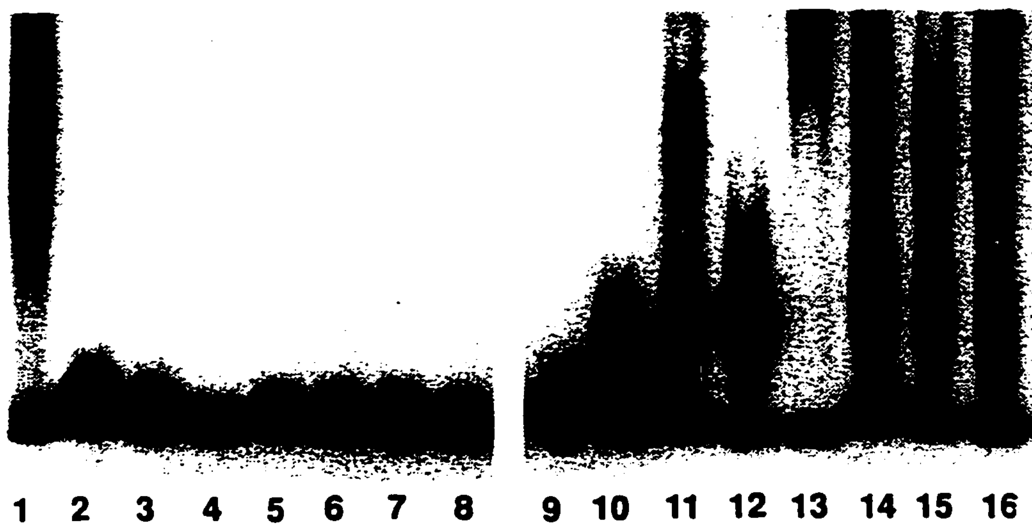
Figure 2-3. Appearance of normal and mutant alleles following agarose gel electrophoresis of PCR products and Southern blotting. The top panel (**A**) in this figure shows the appearance of PCR products (using primers 406 and 409) generated with the genomic DNA of DM and normal individuals following electrophoresis on a 1% agarose gel containing ethidium bromide. The two external lanes correspond to the 123 bp ladder. The sequence of the oligonucleotide PCR primers used in this reaction and their locations are indicated in Figure 2-1. The normal alleles (5-35 CTG repeats) are the bottom bands present in all lanes. Lanes 1 and 10-16 also contain mutant allele bands or smears. Non-DM individuals can appear as homozygous for the normal allele (one band) or as heterozygous (two bands). DM individuals (lanes 1 and 10-16) should appear as one normal size band and an upper faint band or smear that in most cases is not visible in these gels (lanes 11 and 13-16). Samples in **A** correspond to the same as in **B**. (**B**) Mutant alleles can be detected as smears above the normal allele band on Southern blots following the transfer of these products to a nylon membrane and probing with a $(CAG)_{10}$ oligonucleotide probe.

Bottom panel reprinted with permission from Science (Mahadevan et al., Myotonic Dystrophy Mutation: An Unstable CTG Repeat in the 3' Untranslated Region of the Gene, 255: 1253-1255, 6 March 1992) ©1992 AAAS.

A



B



studied revealed its bimodal nature (Figure 2-2). The frequency of some allele sizes is very low and some repeat numbers were not observed. The overall heterozygosity amongst normal individuals in this population was 81%. Analysis of the size of these alleles in different tissues and following intergenerational transmission showed no changes in the number of CTG repeats. This result was obtained through the analysis of the transmission of the normal alleles in over 100 families refered to the laboratory for molecular diagnosis of DM.

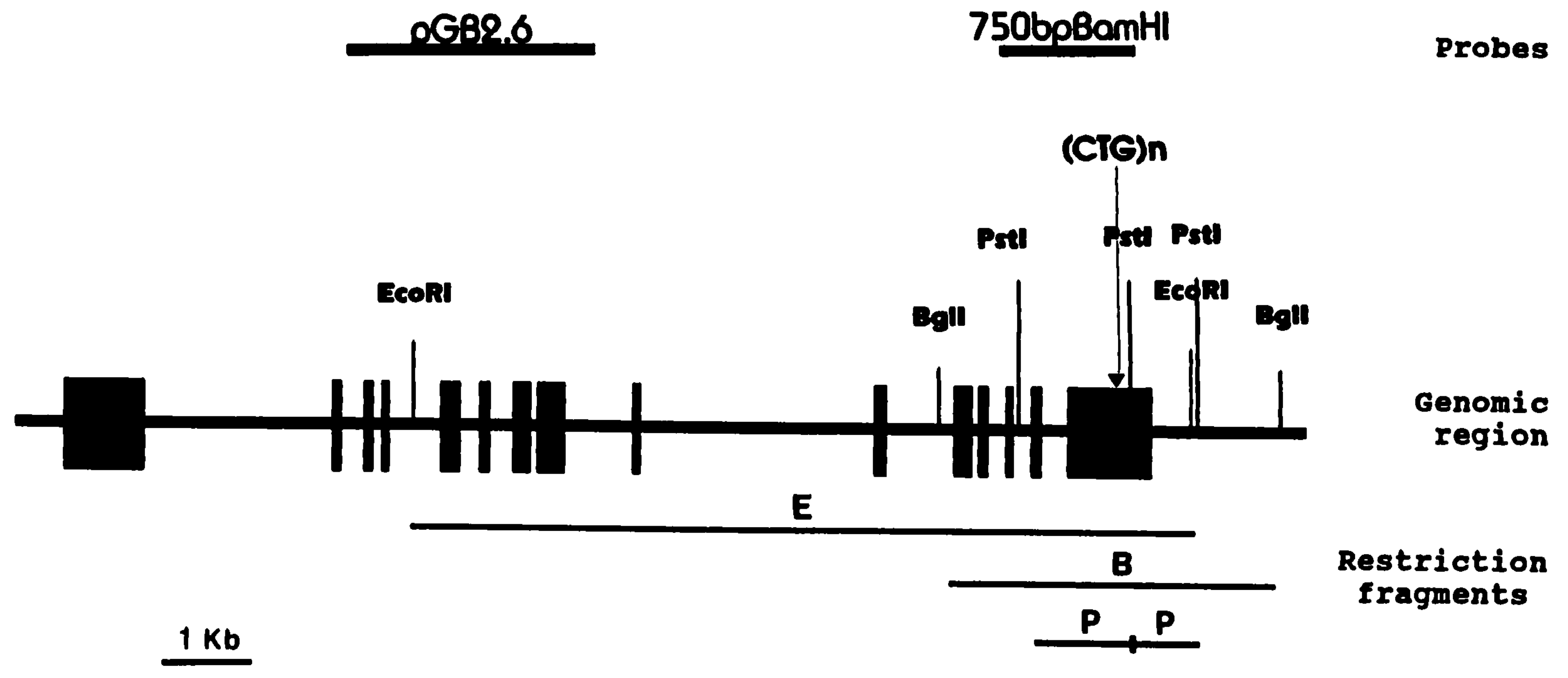
Size distribution and transmission of mutant alleles

Standard PCR and agarose gel electrophoresis could be used for the detection of both normal and mutant alleles. Normal and small mutant alleles (50-200 CTG repeats) could be seen directly on agarose gels stained with ethidium bromide (Figure 2-3 (A)). Southern blotting of these gels and probing with a $(CAG)_{10}$ oligonucleotide probe allowed the detection of mutant alleles in general as bands or smears (Figure 2-3 (B)).

Estimation of the size of the larger mutant alleles was performed by Southern blotting of whole genomic DNA from patients. Restriction digests with enzymes with target sites flanking the mutation allowed the detection of enlarged size bands corresponding to genomic fragments containing the CTG repeat expansion.

Figure 2-4 shows a map of the DM genomic region that includes the DM kinase and the mutation locus in its 3' UTR. The orientation of this map shows the mutation in

Figure 2-4. Map of the genomic region encompassing the CTG repeat in DM. The black boxes represent the 15 exons of the DMK gene. Two probes (**pGB2.2** and **750bpBamHI**), used on genomic Southern blot detection of the level of amplification of the CTG repeat mutation are also shown. The position of the mutation is indicated by a **(CTG)n** within exon 15. The size of three possible restriction fragments is also indicated. The size of the mutant allele (and therefore the number of CTG repeats) can be calculated by subtracting the size of the band of the normal allele from that of the mutant allele in Southern blots. On *EcoRI* restriction digests the mutant allele band also contains the 1 kb insertion/deletion Alu sequence (5.3 kb upstream of the CTG repeat, left) whereas the normal allele presents this segment in only about 50% of the chromosomes.

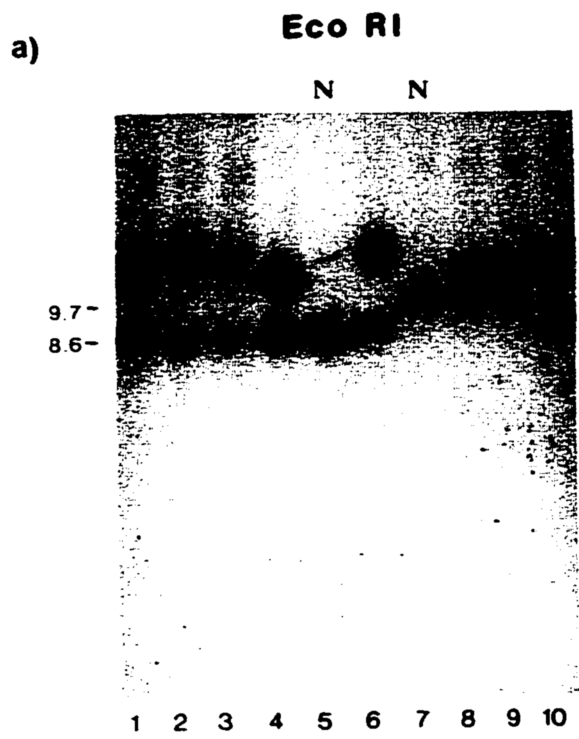

DETECTION OF AMPLIFIED CTG REPEAT SEQUENCES
ON SOUTHERN BLOTS
pGB2.6
750bpBamHI
Probes
(CTG)n
PstI
PstI
PstI
EcoRI
EcoRI
BglI
BglI
Genomic region
E
B
P
P
Restriction fragments
1 Kb

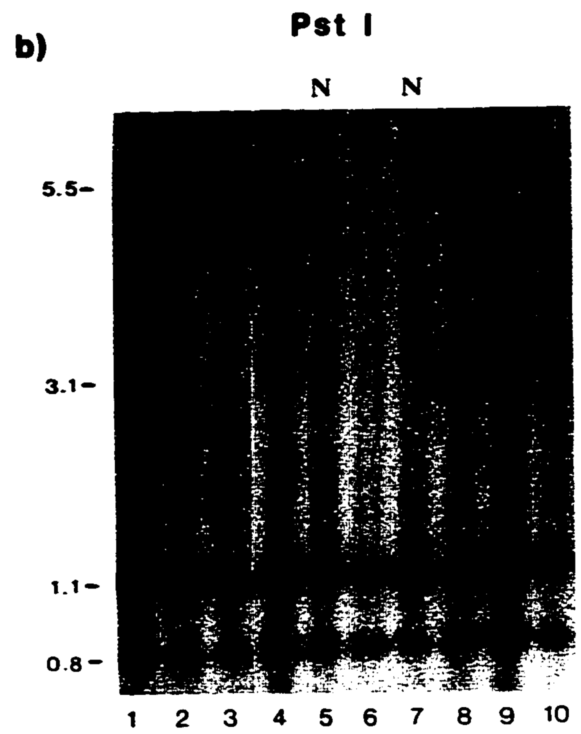
exon 15 of the DMK gene at the centromeric end. The restriction sites and corresponding DNA fragments of three commonly used restriction enzymes are shown. Depending on the enzyme used in restriction digests, it is possible to change the resolution of mutant allele bands. *EcoRI* restriction fragments of the mutant alleles provide higher molecular weight and more compact Southern blot band signals. The probe pGB2.2 maps to this fragment and detects the presence or absence of the 1 kb insertion/deletion polymorphism within it. This 1 kb sequence is in linkage disequilibrium and always present on the same allele as the expanded repeat. *PstI* fragments can provide a better resolution of the size of mutant allele bands than *EcoRI* fragments (Figure 2-5). *BglI* digests provide intermediate size bands. As shown in the map in Figure 2-4, neither *PstI* nor *BglI* restriction digests allow the detection of the 1 kb insertion/deletion sequence 5' to the mutation on Southern blots. Increased mutant allele band size resolution and detection of size heterogeneity is possible with *PstI* restriction digestions. The heterogeneity in the number of repeats present in mutant alleles becomes especially obvious with the use of this enzyme to the point of the apparent disappearance of the mutant allele band in some cases (Figure 2-5 (b), lane 6). Interestingly, the most smeared and faint bands were seen later to correspond to DNA samples of older individuals and of cells kept in culture for a long time period, especially in very large mutant alleles (data not shown).

Study of possible interactions between the normal and mutant alleles

The 5 to 35 repeats present in normal alleles represent another polymorphic marker that

Figure 2-5. Examples of autoradiographs of two possible Southern blotting-probing systems used in order to analyse the level of expansion of the CTG repeat sequence in DM as well as its heterogeneity. The genomic DNA samples obtained from peripheral blood leukocytes on **a)** originate from the same individuals as in **b)** and are in the same numbered order. The **N** above lanes **5** and **7** refers to samples from non-DM individuals and were used as controls. Lane **5** shows the presence of two normal alleles that do not have the 1 kb insertion/deletion Alu sequence and lane **7** shows two normal alleles with this sequence. The numbers beside the lanes refer to the molecular weight of the bands in kb. **a)** Autoradiograph of an *EcoRI* Southern blot probed with pGB2.2. The lower bands (**8.6** and **9.7** kb) correspond to the normal allele. The increased size (over **9.7** kb) of the upper bands reflects the increased number of CTG repeats present in the mutant allele. **b)** Autoradiograph of a *PstI* Southern blot probed with 750bpBamHI. Bands of **0.8** and **1.1** kb correspond to the normal allele. Mutant allele bands or smears are larger. As shown, *PstI* restriction digestions allows a finer detection of size and heterogeneity of mutant allele bands than *EcoRI.* Lane **9** in **b**, for example, shows heterogeneity and mosaicism in the size of the mutant allele band.


a)
Eco RI
N N
9.7
8.6
1 2 3 4 5 6 7 8 9 10


b)
Pst I
N N
5.5
3.1
1.1
0.8
1 2 3 4 5 6 7 8 9 10

could be studied in relation to the level of expansion of the mutant allele. This analysis was performed on over 40 DM individuals selected randomly from over 200 DM pedigrees. The degree of expansion in mutant alleles of those individuals ranged from 50 to over 2000 CTG repeats and the normal alleles had from 5 to 24 repeats. No significant degree of association was found between the sizes of the two alleles (data not shown).

The presence of a series of polymorphic markers within and flanking the DMK gene spanning approximately 15 kb 3' and 15 kb 5' to the CTG repeat (Neville et al., 1994), allowed for a thorough linkage analysis of the DM locus. The segregation of these markers in relation to the expanded CTG repeat in over 100 DM families showed complete linkage disequilibrium. None of the transmissions that we analyzed, except for a rare transmission resulting in an apparent gene conversion (O'Hoy et al., 1993) and described in Chapter IV, gave any indication of recombination events between the homologous chromosomes at this site.

4. Discussion

Sequence analysis of the DM mutation's expanded genomic fragment revealed the presence of a long stretch of CTG repeats at chromosomal location 19q13.3. Standard DNA sequencing and PCR amplification techniques proved to be inefficient on this template and, therefore, techniques were developed in order to detect and quantify the

number of CTG repeats in mutations of different sizes.

Normal alleles had a low number of repeating units, not exceeding 35 in the study group. The transmission of normal alleles both in DM and non-DM families proved to be stable. All the individuals affected with DM or obligate mutation carriers had an enlarged number of repeats associated with the disease. Late onset, mildly affected individuals, start with approximately 50 repeats, but there can be much higher numbers in the rest of the DM patients, sometimes over 2000. This unstable trinucleotide repeat represented a new kind of mutation which was known to occur, at the time of discovery, in only two other genetic diseases: fragile X syndrome (Verkerk et al., 1991; Yu et al., 1991; Oberlé et al., 1991) and spinal and bulbar muscular atrophy (La Spada et al., 1991).

The DM population that was studied originated mainly from the Saguenay-Lac-Saint-Jean region of Quebec. The laboratory had access to the largest collection of peripheral blood leukocyte DNA samples from DM families in the world. This provided the unique opportunity of a thorough study of the behaviour of the mutation on transmission from parent to offspring. The mutation was found to be present in 98% of the individuals with clinical symptoms of DM in the study group (Mahadevan et al., 1992). No other mutations in the DMK gene, or in other genes, were found to be responsible for DM although symptoms of DM were detected in some patients with no DM CTG repeat expansion (Thornton et al., 1994b; Abbruzzese et al., 1996).

Intergenerational transmissions generally resulted in amplification of the mutant CTG repeat array whereby, in most cases, offspring were seen to have larger alleles than their parents, in blood leukocytes. This was seen to occur concurrently with genetic anticipation. The molecular basis of the genetic anticipation observed in DM was then established when various studies showed a positive correlation between the severity of the clinical phenotype and the size of the mutation in DM families (Hunter et al., 1992, Harley et al., 1992b; Tsilfidis et al., 1992).

Although models of triplet repeat sequence instability in human genetic diseases have been proposed (Caskey et al., 1992; Richards and Sutherland, 1992, 1994), the mechanism remains to be established. Amplification of triplets occurs in the absence of recombination between markers flanking the mutation in DM, and the number of repeats in the homologous normal and mutant alleles do not seem to be related. Unlike as suggested for the CAG repeat in Machado-Joseph disease (Igarashi et al., 1996), inter-allelic interaction does not seem to be involved in the intergenerational instability of the DM mutation. The data provide evidence that genetic recombination does not play a major role in the dynamics of the mutation. It may be significant that the mutation rate of tandem-repetitive hypervariable loci in human DNA seems to be proportional to the repeat length (Jeffreys et al., 1988). Long stretches of CTG trinucleotides in DM could allow for slippage during reannealing of the sequence during replication. The proposed model may have to account for intergenerational amplifications of the DM CTG repeat of up to an order of magnitude or more as well

as the occasional size reductions (Chapter III) observed on transmission of mutant DM alleles.

Studies of how different parameters affect the outcome of transmission of the DM mutation may help in the development of an adequate model of its expansion.

CHAPTER III

ANALYSIS OF THE EFFECTS OF THE TRANSMISSION OF THE MUTATION AND ITS INTERGENERATIONAL STABILITY

1. Introduction

This chapter describes studies of the outcome of intergenerational transmission of mutant alleles in DM families as reflected by the number of CTG repeats in the mutation. The emphasis is placed on the transmission of mutant DM alleles with a low number of CTG repeats, and which show little or no change in size following transmission.

My analyses provided evidence of the role of the mutant allele size in its own mutability and therefore the role of this component in the possible mechanism of expansion. Moreover, they also provided molecular evidence for a model that could explain the maintenance of the DM allele in the population.

The incidence of DM may be higher than previously reported, given the variability in penetrance, expression, and age of onset. Some DM individuals have a minimal amplification of the repeat in peripheral blood leukocytes, ranging from 50 to 80 CTG copies. In these cases the disease is very mild and tends to go undetected until

ascertained genetically. We designated this class of the DM mutation a <u>proto</u>mutation because it is the mutation itself at an early stage of intergenerational evolution. We set the upper limit of the protomutation size at 80 repeats, this being the upper threshold of observed stable transmission. DM alleles containing over 80 repeats are defined as full mutations and are seen to expand when transmitted from one generation to the next.

Individuals with the protomutation are comparable to those carrying the <u>pre</u>mutation in fragile X syndrome (Oberlé et al., 1991), a genetic disease which involves the amplification of a CGG trinucleotide repeat (Kremer et al., 1991; Verkerk et al., 1991). Individuals with the fragile X premutation appear clinically asymptomatic (Fu et al., 1991). However, as with DM, transmission of the fragile X mutant allele has been associated with amplification of the trinucleotide repeat leading to severe phenotypes.

We conducted a survey among DM families in order to obtain an estimate of the frequency of the protomutation, to examine its inheritance pattern, and to determine the role of the initial degree of amplification in its intergenerational evolution.

The triggers of CTG repeat amplification in DM were not known and the question of how the length and/or purity of the CTG repeat sequence array might affect the outcome of transmission had still barely been investigated. It was a primary aim of my research to begin to explain these processes.

2. Methodology

Population study group

The DM CTG repeat was examined in 536 DM and 651 non-DM individuals from 158 families. The majority were of French-Canadian origin.

DNA preparation and mutation size analyses

DNA was isolated from peripheral blood samples. Detection of the level of expansion of the DM allele in all of the individuals involved in the study was done by PCR analysis and Southern blotting. Genomic DNA was PCR-amplified using CTG repeat flanking primers (406-409 or 409-410) and a standard protocol. Protomutation allele sizing was performed through PCR amplification followed by ethanol precipitation and resuspension of the PCR product in 1/5 of the original reaction volume. This was followed by 1% agarose or 1% agarose-2% NuSieve (FMC BioProducts) gel electrophoresis against sequenced protomutation alleles and molecular-weight standards. Stable transmission was defined as no apparent differences in electrophoretic mobility (the margin of error was estimated at up to ±3 repeats for protomutation alleles). Sizing of DM alleles exceeding 150 repeats was done by Southern blotting on genomic DNA as described. Alleles containing between 100-150 repeats were sized by PCR analysis followed by Southern blotting of the products as described. In addition to the [^{32}P]-$(CAG)_{10}$ oligonucleotide probe, an alkaline phosphatase-conjugated $(CAG)_{10}$ oligonucleotide probe was used for assessment of the number of CTG repeats. DNA

sequencing revealed that the electrophoretic mobility of the CTG repeat was consistent with the DNA size markers. Due to increased somatic heterogeneity in some large DM alleles, assignment of number of repeats was decided on the basis of an estimated average.

DNA sequencing

PCR reactions were prepared using genomic DNA (as template) from several individuals that had <80 CTG repeats at the DM locus and showed its stable intergenerational transmission. The reactions contained CTG repeat flanking primers 406 and 483 (5'-TGTGGGCCAGTGCCCGCTG-3'), (DMK gene, GeneBank accession number L08835), which flank the repeat sequence within 111 and 648 bp respectively. PCR was conducted as described in the presence of a gene 32 protein (a single-stranded DNA binding protein) (see Appendix). The products were submitted to electrophoresis on a 1% agarose gel and the mutant allele band corresponding to each DM individual was excised and purified with the Qiagen® PCR DNA purification kit according to the manufacturer's instructions. The sequencing reaction used the internal primers 409 or 485 (5-TGCCACCCGCTTAGCTGCG-3') in the forward sequencing reaction and 410 or 520 (5'-GCCCATCCACGTCAGGGC-3') in the reverse sequencing reaction. Purified DNA sequencing was performed on an ABI 373A automated sequencer using Taq polymerase and fluorescent dye labelled dideoxynucleotides, according to the manufacturer's instructions (Applied Biosystems). The sequence data were analysed by ABI software on a Macintosh computer.

2. Results

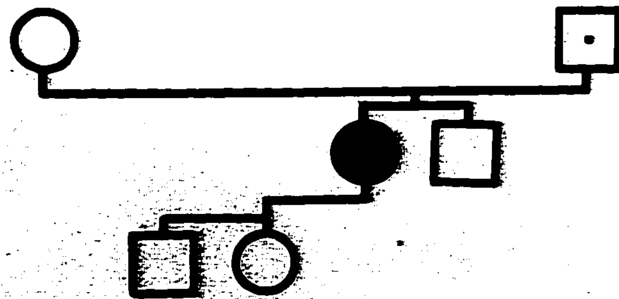
Intergenerational amplification of the DM mutation

Expanded DM alleles were found in the probands who had a clinical diagnosis of this disease. When the mutation was traced back to the previous generation it was found that, in general, a smaller mutant allele could be found in one of the two parents. Figure 3-1 provides an example of three generations of a DM family in which the expanded allele in the proband could be traced back to the father. In this case the father had a minimally expanded mutant allele that was seen to be greatly expanded in the next generation.

Size of the mutation in offspring as a reflection of the size of the mutation in parents

The study of the effect of the size of the mutant allele on transmission outcome revealed an interesting trend. The 320 parent-offspring transmissions analysed were divided into subsets depending on the size of the mutant allele in peripheral blood leukocytes in parents. Overall, the mean intergenerational increase in the number of repeats was a function of the number of repeats present in the mutant allele in those parents. The results could be represented as a sigmoid curve where the smallest parental allele sizes (50-79 CTG repeats) led to relatively small intergenerational amplifications. More pronounced amplifications occurred as the size of the parental allele increased up until a certain parental allele size range (>700 CTG repeats) where intergenerational

Figure 3-1. Transmission of the mutation in a DM kindred. This figure represents three generations of a family in which a small size DM mutation is transmitted from a father that has minimal clinical symptoms of the disease (square with dot, top right) to his daughter (full black circle), who presents DM clinical symptoms. Solid symbol means clinically affected and open symbol not clinically affected. The bands in this figure correspond to a Southern blot of an *EcoRI* restriction digest probed with pGB2.2. Lower and upper bands correspond to the normal alleles (of approximately 9 and 10 kb respectively). In individuals with the mutation the upper band is larger although this may be undetected by genomic Southern blot analysis (see last lane as an example) or much larger (lane that corresponds to the proband, or full black circle) depending on the number of CTG repeats present.



amplifications became more moderate and in some cases intergenerational reductions were seen.

Mutation frequency and protomutation stability

We surveyed the status of the DM locus in 158 families (536 DM mutation carriers, with and without clear clinical symptoms, and 651 non-DM individuals). All the families were ascertained by the clinical diagnosis of the proband; an expanded DM allele was subsequently identified in each of the probands and the rest of the mutation carriers were either first of second degree relatives of those probands. PCR amplification of the repeat sequence, using primers immediately flanking the repeat site (Mahadevan et al., 1992), permitted the detection of small changes in the allele size of normal individuals (5-35 repeats) and in DM mutation carriers with minimal expansion (50-150 repeats). Short length mutant allele bands were seen on agarose gels stained with ethidium bromide (Figure 3-2). Larger lengths (150 repeats or over) were seen as more smeared bands; these could also be seen as hybridization signals on Southern blots of *EcoRI* digested genomic DNA hybridized with probe pGB2.2.

It was found that 11% (60/536) of the DM individuals studied carried the protomutation (50 to 80 repeats). In 10 of the larger DM kindreds the frequency of the protomutation among DM mutation carriers was much higher (55%). The protomutation was found at a very high frequency in the lateral branches of these extended families. In such cases the protomutation was seen to be inherited with either no change or a very small

Figure 3-2. Appearance of the CTG repeat following PCR amplification of the mutation in DM individuals with a low number of repeats in the mutant alleles. This is an example of a 1% agarose gel containing the products of a PCR reaction using primers 406 and 409 and stained with ethidium bromide. The upper band represents DM alleles ranging from 50 (lane **1**) to 150 (lane **8**) repeats. The lower band shows the normal size allele in these individuals. A 123 bp DNA size marker flanks the sample lanes.

Reprinted with permission from Human Molecular Genetics (Barceló et al., Intergenerational stability of the myotonic dystrophy protomutation. 2, 705-709) ©1993, Oxford University Press

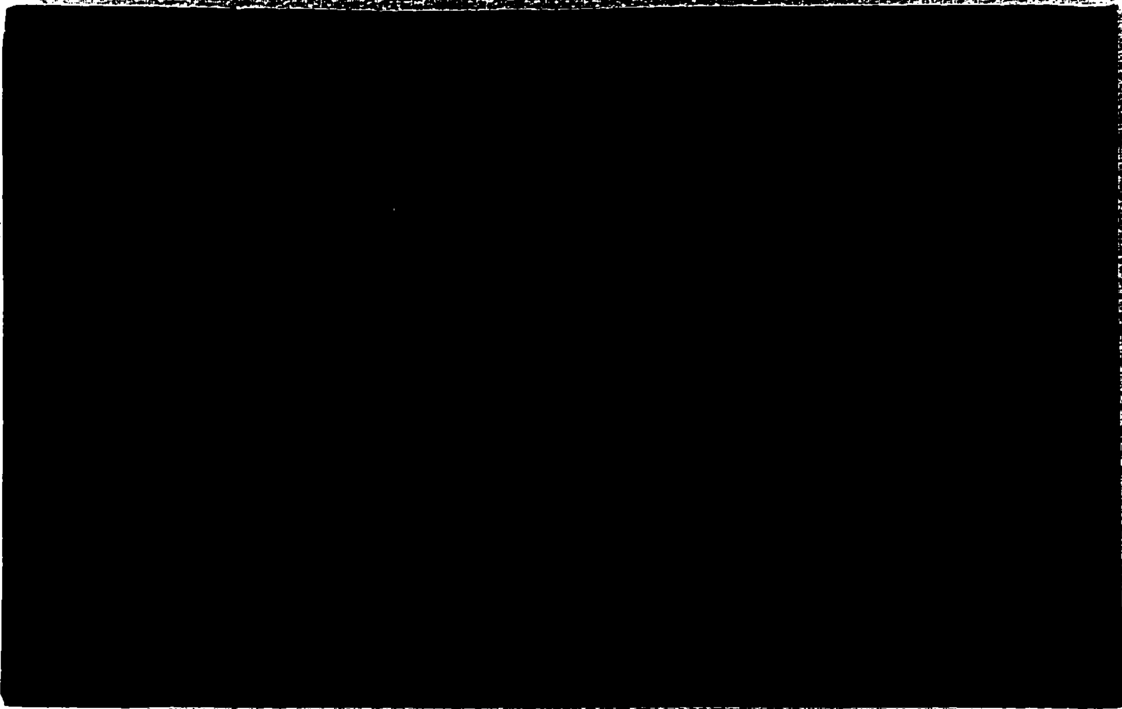

1 2 3 4 5 6 7 8

Figure 3-3. Example of a kindred exhibiting intergenerational stability of the protomutation. PCR products of the DM locus using primers 406 and 409 were submitted to 1% agarose gel electrophoresis, Southern blotting, and probing with an alkaline-phosphatase conjugated $(CAG)_{10}$ oligonucleotide probe. In this kindred the mother (circle with dot) transmits the mutation to five of her offspring. Only one of her daughters (full black circle) inherits a greatly expanded allele. The number of repeats is indicated. Lane **9** contains the maternal protomutation allele. Relative intergenerational stability occurred in four of the five offspring that inherited the mutant DM allele (lanes **2**, **3**, **5** and **6**). Lane **8** shows the smear typically appearing on PCR amplified DM loci containing from 160 to 300 CTG repeats. The shadow bands visible in lanes **2**, **3**, **5**, **6**, and **9** are commonly found in PCR products of protomutation alleles and have been attributed to somatic heterogeneity and/or the subproducts of standard PCR amplification of long sequences of CTG repeats.

Reprinted with permission from Human Molecular Genetics (Barceló et al., Intergenerational stability of the myotonic dystrophy protomutation. 2, 705-709) ©1993, Oxford University Press

13
5
70
11
1 2 3 4 5 6 7 8 9

increase in the number of repeats (Figure 3-3). In one large kindred, the relatively stable transmission of an approximately 60 (±3) repeat allele through four successive generations was observed (data not shown).

Transmission of the DM protomutation

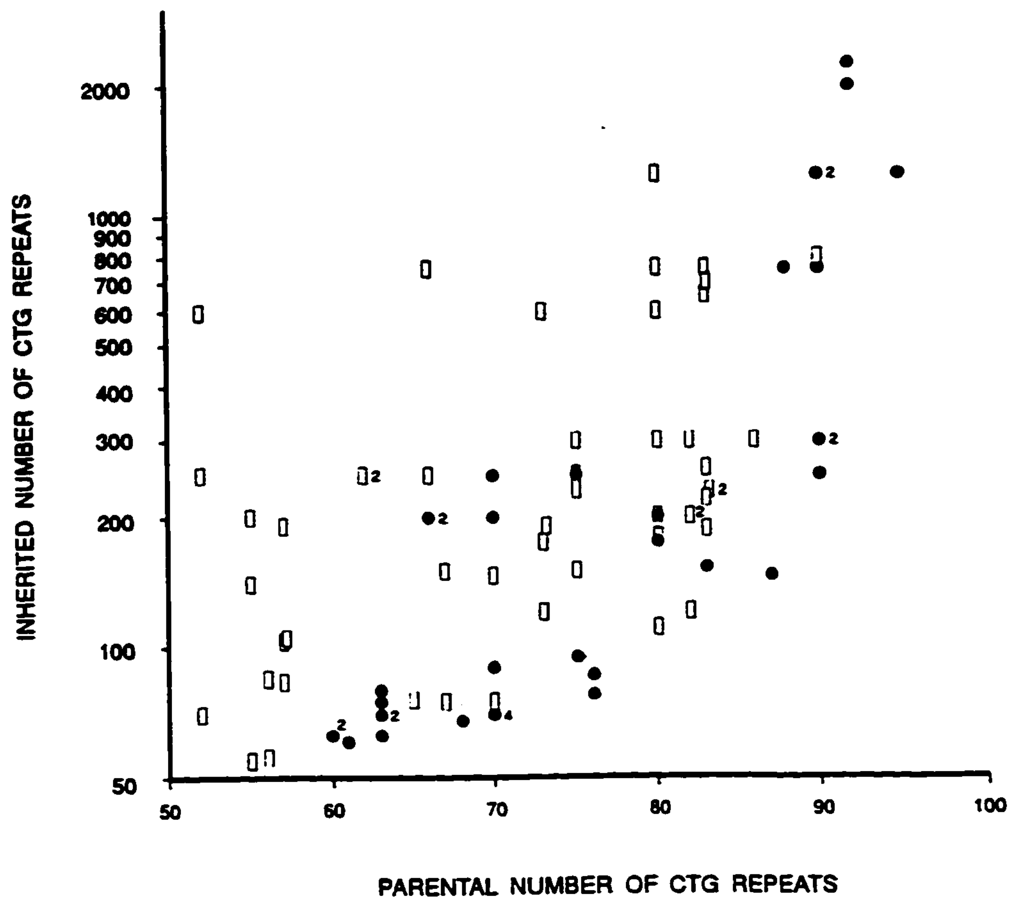
When DNA was available to trace the mutation back for several generations it was found that the protomutation was present in the oldest DM mutation carriers (29 kindreds). No cases were observed in which the protomutation originated from a normal size allele. The lowest number of repeats observed in a DM allele in the study group was 50.

Sixty DM-parent/DM-offspring pairs in which the parent had the DM protomutation were identified; 20 resulted in the inheritance of the protomutation (50-80 repeats), that is, no change (n=11) or a small size increase (n=9) was observed. No reductions of the DM allele as a result of transmission of a protomutation were observed.

Some gender differences in the intergenerational amplification of the DM protomutation can be seen (Figure 3-4). The majority of parents that were DM protomutation carriers were males (36/60). Only eight paternal transmissions resulted in the inheritance of a protomutation. In contrast, 16 maternal transmissions resulted in the inheritance of a protomutation. Only transmissions from protomutation carrying fathers resulted in alleles containing over 600 repeats in this study group and protomutation carrying

Figure 3-4. Number of repeats inherited from DM parents with mutant allele sizes ranging from 50 to 100 repeats. A rectangle represents a transmission from a male and a circle represents a transmission from a female. The number beside some of the symbols represents the number of overlapping data points at the same position. The graph shows the general trend of diminished intergenerational stability as the number of CTG repeats in the parental allele increases. Small protomutations can evolve into full mutations in a single step and the frequency of this occurrence increases with the increasing size of the parental allele.

Reprinted with permission from Human Molecular Genetics (Barceló et al., Intergenerational stability of the myotonic dystrophy protomutation. 2, 705-709) ©1993, Oxford University Press



INHERITED NUMBER OF CTG REPEATS
2000
1000
900
800
700
600
500
400
300
200
100
50
50
60
70
80
90
100
PARENTAL NUMBER OF CTG REPEATS

mothers had offspring with a lower number of repeats (Figure 3-4). Further analysis of the size of the DM allele in the progeny of DM protomutation carriers revealed that smaller size alleles were generally more stable than larger ones. For example, inherited mutant alleles having in excess of 200 CTG repeats were observed in 21% of transmissions from the 50-60 repeat parental group, 32% of transmissions with 61-70 repeats, 57% of transmissions of DM alleles with 71-80 repeats, and 71% of transmissions of DM alleles with 81-90 repeats (Table 3-1). All transmissions of DM alleles with approximately 81 or more repeats resulted in large amplifications, ranging from 120 to 1250 repeats. Amplification of the DM allele did not necessarily occur in all of the DM offspring of a given protomutation carrier and when it did, it was not necessarily to the same extent.

The stability of the DM protomutation is not due to interruptions of perfect CTG repeat arrays

We analysed the possibility that intergenerational stability of the protomutation was due to perfect repeat sequence interruptions. For this purpose we studied the sequence composition of the protomutation in eight DNA samples obtained from patients in which stable transmission had been observed. The sequence of the two (CTG and CAG) complementary strands showed no interruptions or base substitutions (data not shown).

Table 3-1. Summary of amplification results over a range of parental allele sizes

Parental number of repeats	50-60	61-70	71-80	81-90
Number of DM-parent/DM-offspring pairings	14	25	21	24
Number of cases with stable transmission	3 (21.5%)	7 (28%)	1 (5%)	0 (0%)
Number of cases with allele size under 200 repeats	8 (57%)	10 (40%)	8 (38%)	7 (29%)
Number of cases with allele size over 200 repeats	3 (21.5%)	8 (32%)	12 (57%)	17 (71%)

Reprinted with permission from Human Molecular Genetics (Barceló et al., Intergenerational stability of the myotonic dystrophy protomutation. 2, 705-709) ©1993, Oxford University Press

4. Discussion

This chapter deals with the overall study of DM mutation size and composition as a relevant component of intergenerational dynamics. The relatively high number of transmissions analyzed and the overall consistency of the results obtained give validity to such a study. However, since data on mutation size was obtained from only one of the tissues in those patients (peripheral blood leukocytes), including individuals of all ages, there is a limit to the extent to which correlations could be made. The results of these analyses indicated, nevertheless, that *the size of the mutation must be an important component of the intergenerational dynamics of the CTG repeat array.*

Molecular analysis of DM kindreds is typically only done following ascertainment of a symptomatic individual. None of our kindreds was ascertained through a proband with a protomutation. Nevertheless, over 11% of individuals with the DM mutation in our survey (60/536 DM mutation carriers with and without clear clinical symptoms of the disease) had the protomutation. Furthermore, the frequency of the protomutation among DM mutation carriers was considerably higher than 11% in many of our larger DM kindreds, and the frequency was highest in the lateral branches of these extended families. Because the data analysed in this study originate from families ascertained with DM, where amplification of a protomutation had already occurred, they are not truly representative of the situation in the general population. Only a very large population screening, such as that of Reiss et al., (1994), or Rousseau et al., (1995), on

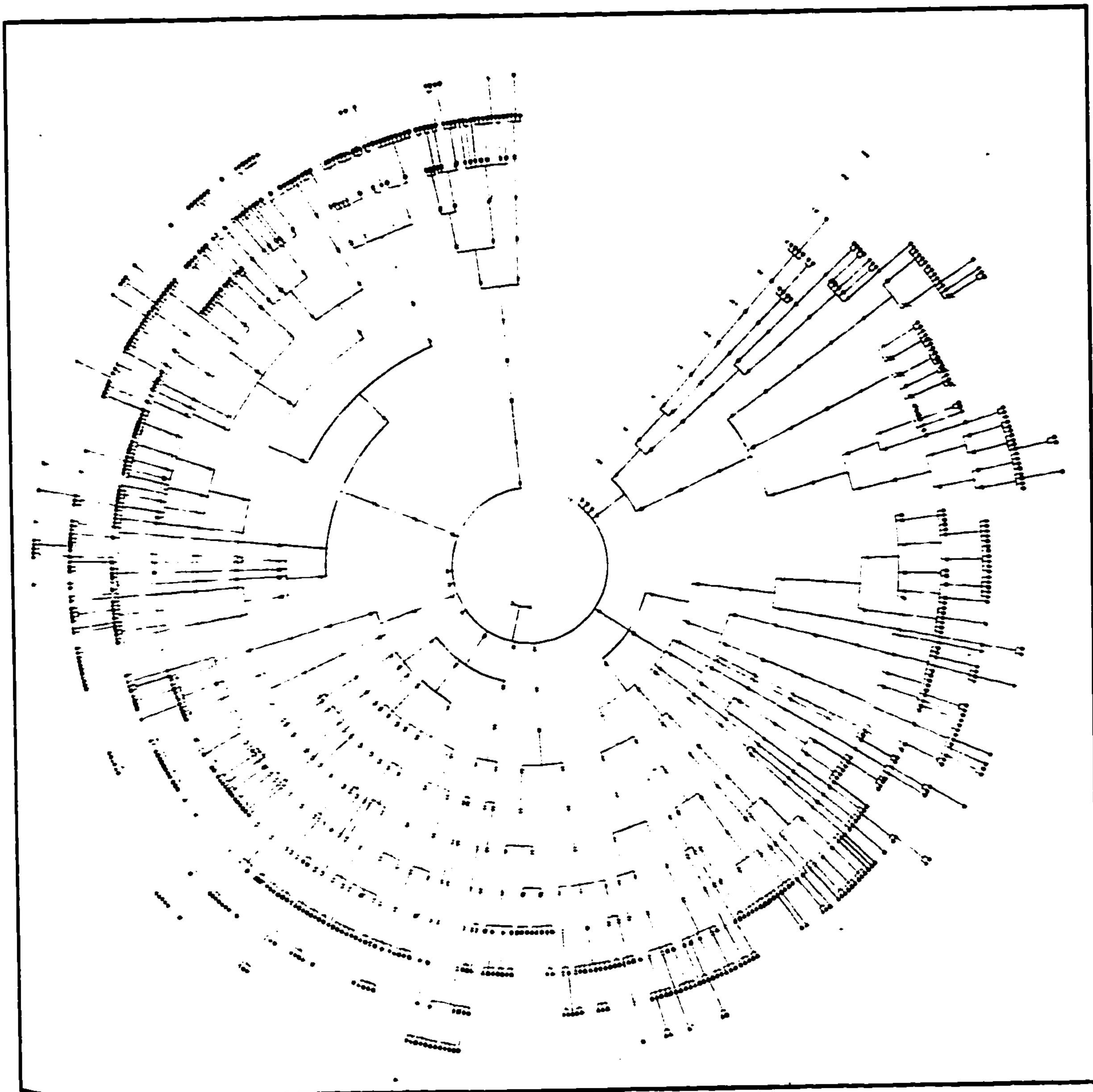
the prevalence of the premutation in fragile X syndrome, would provide reliable indices of DM protomutation prevalence and stability. What this study suggests, nevertheless, is a *higher frequency of the DM allele than previously estimated*, as well as a high level of stability of the DM protomutation in the general population.

It was clear that the DM protomutation could be transmitted from one generation to the next with minimal or no amplification. In the study group of 158 families, 20 of the 60 DM-protomutation-parent/DM-offspring pairings were characterized by the inheritance of a protomutation (that is, the offspring inherited a maximum of 80 repeats). Similar to the inheritance observed for the fragile X premutation, we observed the stable transmission of the DM protomutation through four generations in one large kindred. *These results suggest a model in which the maintenance of this disease allele in the population is based on the existence of a metastable protomutation with sufficient stability to be passed down through many generations.* This could explain the persistence of an autosomal dominant disease such as DM, despite the presence of genetic anticipation and low reproductive fitness of severe DM phenotypes.

A genealogical reconstruction of DM in the Saguenay-Lac-Saint-Jean area (Quebec, Canada), (Mathieu et al., 1990), where most of the DM families originated, revealed that all patients in that study could be traced back to a couple who settled in "Nouvelle-France" in 1657 (Figure 3-5). This genealogical reconstruction provides a strong argument in favour of the genetic homogeneity or common origin of

Figure 3-5. Genealogical reconstruction of DM in the Saguenay-Lac-Saint-Jean region. DM could be traced back to one couple (centre of the pedigree) that immigrated to this region from France in the seventeenth century. This pedigree represents 746 individuals over 14 generations. From generation II to generation IX only the individuals responsible for the kinships are illustrated. In the more recent generations, the asymptomatic healthy persons were omitted. Square = male; circle = female; / = deceased; solid symbols = affected; half-solid symbols = carrier of a partial syndrome; grey symbols = screened.

Reprinted with permission from Neurology (Mathieu et al., Genealogical reconstruction of myotonic dystrophy in the Saguenay-Lac-Saint-Jean area (Quebec, Canada) 40, 838-842) ©1990, Little, Brown and Company Publishers



the mutation, at least within the French-Canadian population. This, together with the total linkage disequilibrium observed between the DM mutation and nearby polymorphic markers (Mahadevan et al., 1993a; Neville et al., 1993; Whiting et al., 1995b), again lends credence to the hypothesis that originally only one or at most very few amplifications of normal size alleles resulted in the first protomutation(s).

Only one case of a recent formation of a DM mutation has been reported (Krahe et al., 1995b). Since such events seem very rare, the maintenance of a relatively high prevalence of this disease allele in the population, despite the rapid loss of reproductive fitness in affected families, could be explained with the maintenance of an intergenerationally stable form with no or hardly any clinical consequences. The number of progeny from DM protomutation carriers in our study group gives no indication of diminished reproductive fitness.

The maintenance of this disease allele in the population could also be explained by reversions of the amplification process, that is, by intergenerational reductions of the number of CTG repeats, as reported by our laboratory (O'Hoy et al., 1993) and others (Shelbourne et al., 1993; Brunner et al., 1993a; Ashizawa et al., 1994). Reductions resulting in alleles containing over 80 repeats could prolong the reproductive fitness of DM kindreds (perhaps by one or two generations). However, only the reduction events leading to allele sizes in the range of 50 to 80 repeats, would contribute to the pool of protomutation carriers and therefore to the possible long-term transmission of these

disease alleles. Even if such reductions were a source of protomutations, the long-term maintenance of the DM allele in the population would still depend upon the intergenerational stability of this allele class.

A variety of gender differences in the transmission of DM protomutations were observed in the study group. The majority of DM protomutation carriers were males and no mothers with less than 60 repeats were detected. Intergenerational amplifications from fathers led to larger allele sizes (Figure 3-4). A similar phenomenon was observed when the study of the transmission of DM mutations with less than 100 CTG repeats in 38 Dutch kindreds revealed an excess of males who did not manifest disease. These males preceded the generation of the probands, and therefore individuals with considerably expanded mutations (Brunner et al., 1993b).

In fragile X an overlap was found between the highest number of CTG repeats in normal individuals and the lowest number of the same repeats in individuals from these families (Fu et al., 1991). In DM such overlaps were not observed. Nevertheless, as proposed for the CGG repeat in fragile X (Fu et al., 1991; Yu et al., 1992), a direct relationship between increased DM trinucleotide repeat number and repeat instability was observed. Amplification of a given protomutation occurred with a frequency that was influenced by the number of repeats present (Figure 3-4 and Table 3-1). Transmission of DM alleles containing 80 or more CTG repeats resulted in significant amplification. The strong correlation between allele length and the frequency and extent

of repeat sequence amplification provides a clear indication of the *pivotal role of the DM CTG repeat in its own amplification process*. While the data presented strongly suggest that instability of the DM locus is primarily due to its increased size, the absence of a complete allele-length/allele-amplification correlation suggests a *role for other as yet unidentified factors*.

Some size ranges appear more frequently than others. An example of this is the absence of any DM protomutation carrier offspring bearing repeats in the 300-600 repeat range (Figure 3-4). This may reflect an extreme instability of this size range, the underlying amplification process, or it may be merely a sampling artefact.

In other genetic diseases with a similar molecular basis it was found that interruptions (such as base substitutions) of the perfect repeat sequence array led to an increase in allelic stability (Eichler et al., 1994; Chung et al., 1995). Sequence analysis of some of the most stable protomutations revealed no interruptions.

The fact that instability is dictated by a minimum threshold of CTG repeats suggests a link between a novel structure acquired when the number of repeating units is sufficiently high and the mechanism of expansion of this repeating unit. The fact that instability increases dramatically beyond the upper threshold of protomutation size may suggest a sequence-structure-linked mechanism of expansion. The perfect repeat sequence interruptions found in the more stable mutant alleles in other trinucleotide

repeat diseases should perhaps be expected to interfere with this structure and mediate changes in the dynamics of those mutations.

The results of this study also suggest that the gender of the parent affects the intergenerational dynamics of the DM mutation. This is examined in the next chapter.

CHAPTER IV

PARENT OF ORIGIN DIFFERENCES IN EXPANSION OF THE DM MUTATION AND THE CASE OF CONGENITAL DM (CDM)

1. Introduction

Previous studies (Chapter III) suggested that the gender of the transmitting parent, at least for mutant DM alleles containing <80 CTG repeats, had an effect on the intergenerational evolution of this mutation. This chapter analyses in more depth the possible differential effect of maternal and paternal transmissions both on the intergenerational dynamics of the CTG repeat and on disease severity.

I studied the size of the mutation in males and females in general as well as the differences in outcome when the CTG repeat is transmitted by male and female parents. Different trends were seen in both cases. Fathers with a high number of repeats showed a higher probability than mothers of intergenerational reductions of the mutation. So, in general, the outcome of transmission was seen to be dependent on a combination of the size of the mutation and the gender of the transmitting parent.

Another gender-related component of this mutation plays a role in CDM. Although both males and females seem to have the same risk of inheriting this form of the disease, the

only unquestionable cases of CDM were seen following maternal transmission of the mutation. I conducted a study of the maternal and paternal transmission of the mutation and compared the transmissions that resulted in CDM to those that resulted in the adult form of the disease. The mothers and their CDM infants were shown to have CTG repeat numbers in mutant alleles which were, on average, significantly higher than those seen in non-CDM (Tsilfidis et al., 1992; Barceló et al., 1994).

The identification of the factors responsible for the gender-related differences in transmission should help to focus the research towards the understanding of the mechanism(s) of expansion of this mutation.

2. Methodology

Population study group

Analysis of mutation status for the purpose of examining the differential effects of maternal and paternal transmission was done on a total of 174 DM-mother/DM-offspring and 126 DM-father/DM-offspring pairs. CDM studies were done on over 300 transmissions of the mutation (174 maternal and 127 paternal). This included 37 CDM cases and their mothers.

Myotonic dystrophy diagnosis

The clinical diagnosis of DM was as described previously (MacKenzie et at., 1989)

adhering to published recommendations (Griggs and Wood, 1989). The diagnosis criteria for congenital DM were: (1) marked hypotonia at birth; (2) respiratory failure necessitating ventilation at birth; (3) absence of an underlying condition which would lead to hypotonia or respiratory failure (for example, infection); and (4) a diagnosis of DM in the infant's mother. Additional signs associated with the disorder, including talipes, facial displegia, the presence of polyhydramnios and reduced fetal movements were noted. However, the latter findings alone, in the absence of the above criteria were not sufficient for the diagnosis of CDM. Only 47% of the CDM infants in one study (Harper, 1975) had respiratory failure. It is therefore likely that infants with milder forms of CDM existed in the population but were not diagnosed. Nevertheless it is likely that the ascertainment of the more clinically relevant severe CDM approached 100%.

DNA preparation and analysis

Genomic DNA (2-5μg) was digested with *EcoRI* (NEB), submitted to electrophoresis on 0.8% agarose gels, and transferred onto nylon membranes. Southern blots were probed with pGB2.2. Blots were washed in 0.2 x SSC and 0.1% SDS at 65°C and exposed to X-AR autoradiographic film for 1-4 days. CTG repeat array length in mutant alleles was assigned based on the mobility of the mutant compared to the normal allele band and a DNA band size marker.

3. Results

General parental-gender dependent effects on the transmission of the mutation

An analysis of the size of the mutation in peripheral blood leukocytes was conducted in over 100 DM families. These families had been refered to the laboratory for molecular diagnosis of the disease. They were ascertained through a member with clinical symptoms of the disease, refered to as a proband. Probands belonged, in the great majority of cases, to the youngest generation in those pedigrees. The analysis of the number of CTG repeats present in the mutant allele in the youngest generation (proband and siblings) in our DM families revealed that the mean number was approximately 800. When these individuals were divided into males and females we found that males presented approximately 770 and females 820 CTG repeats. This difference was not found to be significant. In contrast when, within the same population, only the DM parents of these offspring were selected, on average, the mean number of CTG repeats in the mutant alleles of DM mothers was approximately 350 whereas in fathers it was approximately 230 (Figure 4-1). In addition to having, in general, a higher number of repeats in mutant alleles, DM mothers have DM offspring with a significantly larger number of repeats (mean = 930 CTG repeats, 166 chromosomes analysed) than DM fathers (mean = 800 CTG repeats, 124 chromosomes analysed). It may be significant that father to female offspring and father to male offspring transmissions show discordant parental numbers of repeats (174 versus 289 respectively). Moreover, female offspring of those fathers show a mean increase of

Figure 4-1. Maternal and paternal inheritance of the mutation. Mean number of CTG repeats in mutant DM alleles in leukocytes of mothers and fathers and mean increase (offspring minus parent) inherited by their DM male and female offspring (white bars). The numbers on top of the bars refer to the average number of repeats and N is the number of transmissions in each set. Unlike fathers, DM mothers have and transmit, on average, large expansions and show similar patterns of transmission outcome for male and female offspring.

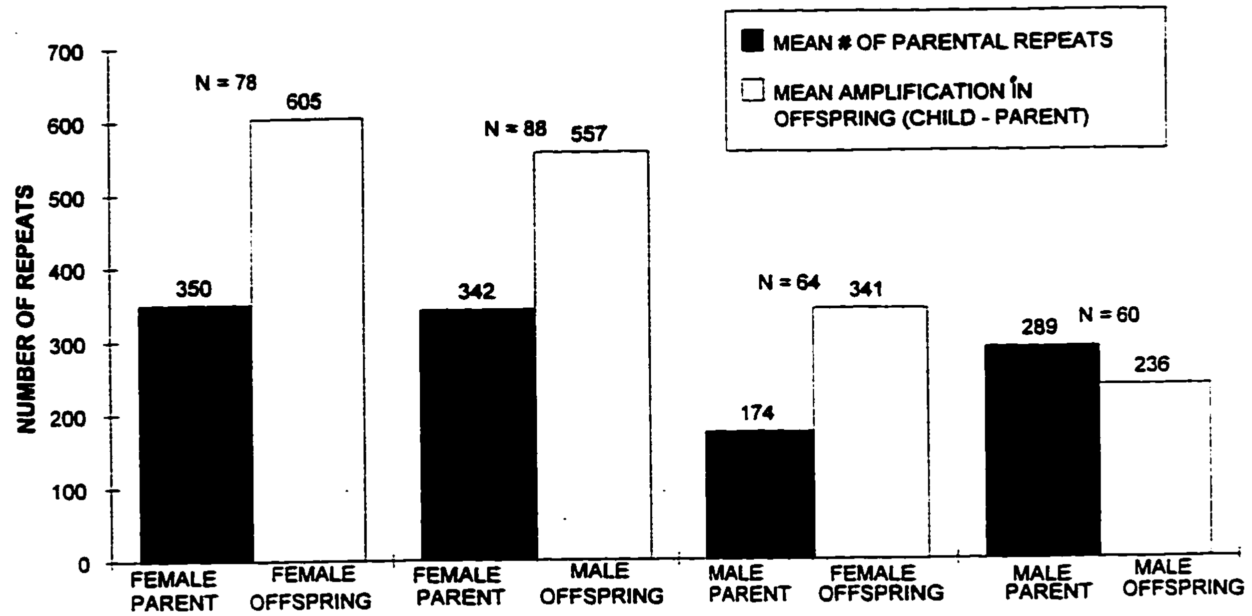
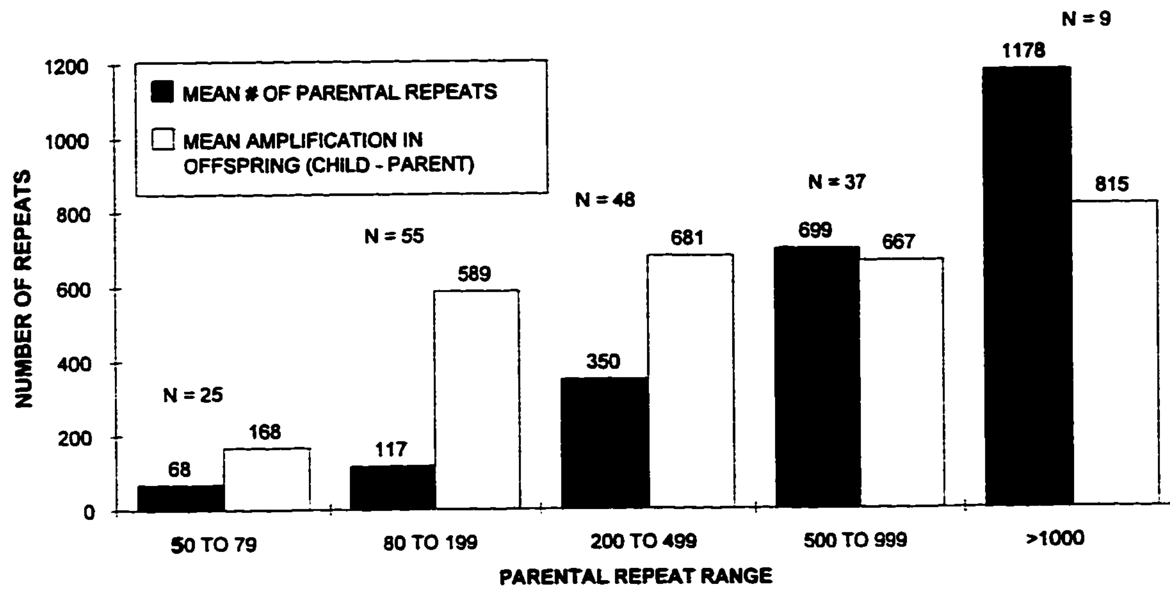
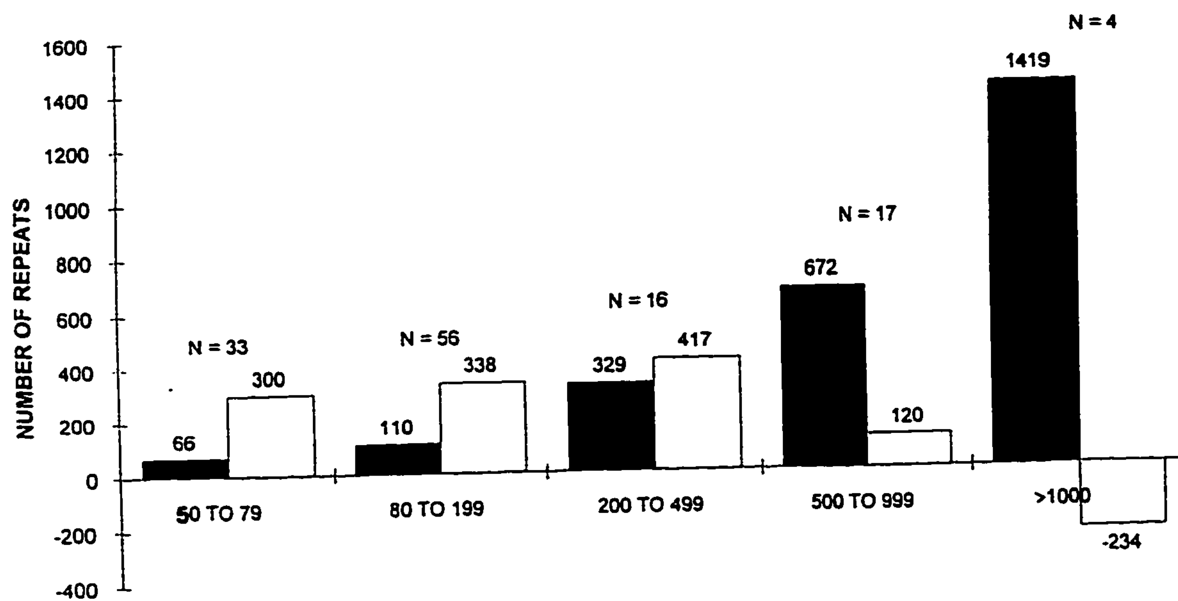

PARENTAL REPEATS TO AMPLIFICATION BY SEX OF PARENT AND SEX OF OFFSPRING
MEAN # OF PARENTAL REPEATS
MEAN AMPLIFICATION IN OFFSPRING (CHILD - PARENT)
NUMBER OF REPEATS
700
600
500
400
300
200
100
0
N = 78
350
605
N = 88
342
557
N = 64
174
341
289
N = 60
236
FEMALE PARENT
FEMALE OFFSPRING
FEMALE PARENT
MALE OFFSPRING
MALE PARENT
FEMALE OFFSPRING
MALE PARENT
MALE OFFSPRING

Figure 4-2. Mean number of repeats present in DM mothers and fathers and increase over this number acquired by their DM offspring. DM parents were subdivided into sets according to number of CTG repeats present in the mutant allele (black bars). The adjacent white bars represent the increase in number (not the total number) of CTG repeats acquired by the mutant allele in their DM offspring. **N** represents the number of transmissions in each set. **A** = maternal transmissions, **B** = paternal transmissions.

A



B



approximately 340 CTG repeats whereas in male offspring this increase is of approximately of 230 CTG repeats instead.

The paternal and maternal transmissions were then segregated according to the number of CTG repeats in the mutant allele in parents (Figure 4-2). The results revealed significantly different patterns of transmission. As the number of CTG repeats in mothers increases from 50 to over 1000, the increase in repeats in offspring shows a slow start, followed by relatively steady intergenerational amplifications that seem to reach a plateau in transmission from parents with a high number of repeats (Figure 4-2 (A)). Paternal transmissions, in contrast, show a greater intergenerational amplification at a low number of repeats that reaches an earlier plateau and overall negative amplification at the high end (Figure 4-2 (B)).

Reductions in the size of the mutant allele on transmission.

Although intergenerational transmission of the DM mutation tends to result in its expansion, four cases of intergenerational reductions were observed in our study group. One of those resulted in the complete loss of the mutation (O'Hoy et al., 1993). In this case, genomic Southern blot and PCR based molecular diagnosis methods revealed that the transmission of a mutant allele from a DM father who had an average of 750 CTG repeats in blood leukocytes resulted in a 13 CTG repeat allele in his normal phenotype daughter. Extensive and high resolution haplotype analysis revealed that this was very likely the result of a recombination event that resulted in discontinuous gene conversion

tracts that included the paternal 13 repeat normal allele sequence. Gene conversion can only be determined by examination of all meiotic products which were not available in this case.

Three other cases of intergenerational reductions in the number of CTG repeats were detected. These losses were more moderate and it did not seem likely that they involved gene conversion events. In the first case the transmitted DM allele of the affected father (with an average of 1000 CTG repeats in the mutant allele of peripheral blood leukocytes) resulted in an affected son with approximately 500 CTG repeats in leukocytes. In the second case a father with approximately 600 CTG repeats in the mutant allele in leukocytes, had a daughter in which the same allele had lost approximately 170 CTG repeats. The only maternal transmission, in the same study group, that resulted in loss of CTG repeats was seen when an individual with approximately 1000 CTG repeats in the mutant allele transmitted the same allele to her son who had approximately 900 repeats. Other, more moderate, losses of CTG repeats on transmission could likely occur and nevertheless be undetected by the currently used molecular diagnosis systems.

Maternal effects in the transmission of the mutation: correlation between CTG repeat length in offspring and frequency of CDM

An increased number of CTG repeats in the mutant allele was found in most CDM individuals as compared to those that had the adult form of the disease. Within non-

CDM individuals the majority had undergone only moderate mutation amplification (less than 3% had over 1500 CTG repeats) and 57% had less than 500 CTG repeats. In contrast, 55% of individuals affected with CDM had over 1500 CTG repeats in their peripheral blood leukocytes. These data suggest a correlation between the CTG repeat length and CDM. Interestingly, a proportion of severe CDM children (5/37) had only modest size mutations (500-1000 CTG repeats), and some infants with non-CDM presented with the largest expansions (over 2000 CTG repeats). Therefore, while amplification of the mutation was seen in all CDM cases, it did seem to be neither a necessary nor a sufficient condition for the development of this form of the disease.

Discordant CTG repeat lengths in the mothers of CDM infants were also found. The majority of DM mothers with non-CDM children had fewer than 500 CTG repeats while mothers of CDM patients tended to show larger expansions. The same study showed that the degree of DM-mother/DM-offspring intergenerational amplification does not appear to be greater in CDM than in non-CDM (Tsilfidis et al., 1992).

Additive influence of maternal and offspring DMK gene CTG repeat lengths in the genesis of CDM

I investigated the possibility that CDM occurs only following maternal transmission of the mutation due to the fact that fathers do not transmit large enough alleles to their offspring. Analysis of the dynamics of the mutation in over 300 transmissions (174 maternal and 127 paternal) indicated that while a high number of repeats seems to be

a necessary condition for CDM it cannot, by itself, explain its exclusive maternal inheritance. One-quarter of the offspring in this group who inherited the mutant allele from their fathers had CTG repeat arrays larger than those found in the CDM cases with the lowest number of CTG repeats (around 700 repeats).

CDM offspring from mothers with a relatively low number of repeats generally had a number of repeats in the upper end of the CDM range. Mothers with a relatively high number of repeats tended to have a higher percentage of CDM offspring, even when the number of repeats in these offspring was in general at the lower end of the CDM spectrum (Figure 4-3).

4. Discussion

I conducted an overall survey of gender-related differences that might affect the dynamics of the intergenerational transmission of the DM mutation. In agreement with previous reports (Brunner et al., 1993b; Lavedan et al., 1993a; Harley et al., 1993), the data indicate that *the intergenerational dynamics of this mutation are both linked to the number of CTG repeats and the gender of the transmitting parent.*

Unlike in fragile X (Rousseau et al., 1994; Loesch et al., 1995), males and females with the disease seemed to have the same probability of inheriting mutant alleles of any size. On average, the later generation of offspring in the study group was seen to have

Figure 4-3. Relationship between maternal and offspring numbers of repeats and CDM. Squares represent the transmissions resulting in the congenital form and circles represent the transmissions resulting in the adult form of DM. The distribution of these symbols indicates a tendency for CDM to occur when both the maternal and the inherited number of repeats in blood leukocytes is high.

Reproduced with permission from The American Journal of Human Genetics (Barceló et al., Additive Influence of Maternal and Offspring DM-Kinase Gene CTG Repeat Lengths in the Genesis of Congenital Myotonic Dystrophy. 54, 1124-1125) ©1994, The University of Chicago Press

4000
3000
2000
1000
0
OFFSPRING NUMBER OF REPEATS

0
500
1000
1500
MATERNAL NUMBER OF REPEATS

■ CONGENITAL
○ NONCONGENITAL

practically the same number of repeats regardless of the gender. Although the data presented in Figure 4-1 for male transmissions of the mutation may not give a clear indication of this fact, the explanation for the discordance in the size of the mutation in male and female offspring of DM fathers could be found when parental transmissions were divided according to the number of repeats in males and females (Figure 4-2). Nevertheless, as seen in the fragile X syndrome (Fish et al., 1995), SBMA (Biancalana et al., 1992), HD (Kremer et al., 1995), SCA1 (Jodice et al., 1994), and other trinucleotide repeat mutation diseases, the gender of the parent had a clear effect on the intergenerational dynamics of the CTG repeat in this disease. This is reflected in the fact that changes in the CTG repeat happen to different degrees depending on which parent (mother or father) with a defined size allele transmits the mutation. These differences are most striking in transmissions from parents with a high number of CTG repeats where, unlike for mothers, paternal transmissions result, in a high percentage of cases, in losses of CTG repeat array tracts. In agreement with our results, an international research report (Ashizawa et al., 1994), that grouped together the intergenerational transmissions that resulted in reductions, revealed that paternal transmissions result in a higher and more dramatic frequency of reductions (loss of repeats) than maternal transmissions. Those reductions were seen to occur at a higher frequency from fathers that had a high number of repeats in the mutant allele in blood leukocytes. The reason(s) behind the differences in the size of the inherited mutant allele in maternal and paternal transmissions are not known.

Inheritance of CDM from mothers

While the inheritance of the different allele sizes may not be linked to gender, DM mothers have a significantly higher number of CTG repeats than fathers (Figure 4-1). The reason for this difference seems to have a sociological basis.

The CDM offspring of these mothers also had a higher number of CTG repeats in mutant alleles, although intergenerational amplification of the mutation from those mothers did not appear to be greater in CDM than in non-CDM.

Reports on the observation of a higher number of CTG repeats in mutant alleles of CDM patients (Tsilfidis et al., 1992) were combined with the observation of CTG repeat losses on transmission of large size mutant alleles from fathers in small population study groups. These led some investigators (Mulley et al., 1993; Lavedan et al., 1993b) to propose that the exclusive maternal inheritance of DM was derived from the incapacity of fathers to transmit large enough alleles for CDM to occur. My study of this possibility revealed that a considerable number of paternal transmissions resulted in mutant alleles in offspring that were as large or larger than those seen in CDM. Therefore, the number of repeats in those offspring does not seem to be a sufficient condition for CDM to occur. It was also noted that the number of repeats in CDM offspring of DM mothers did not need to be very high if those mothers had a relatively high number of repeats themselves. Hence, *a very high number of repeats in offspring is neither a necessary nor a sufficient condition for CDM to occur.*

These two observations, together with previous studies on maternal phenotypes of CDM offspring (Koch et al., 1991), suggested that CDM arises from an "additive" pathology comprised of both maternal and offspring components (Barceló et al., 1994). This unusual basis for the more severe phenotypic variation of DM could perhaps be included in a broad definition of "imprinting", which has been defined as differential phenotypic manifestations of disease based on the gender of the transmitting parent. Imprinting has been seen to be sex-linked to methylation status differences in CpG sequences that affect the expression of relevant genes. Analysis of the methylation status of DMK and surrounding regions have revealed, so far, no disease- or gender-related differences (Jansen et al., 1993; Shaw et al., 1993).

The differential phenotypic manifestations of CDM may be due to metabolic disturbances of a DM mother on a developing DM fetus with a sufficiently high number of CTG repeats. The DM mother must make some contribution to the phenotype of the CDM fetus in addition to having transmitted a large mutant allele -- since DM fathers can transmit large alleles without giving rise to CDM. The maternal influence may be related to (or mediated by) her phenotype (which is in turn due to the number of CTG repeats in her mutant allele) since only the more severely affected mothers tend to have CDM offspring. A sufficiently high number of CTG repeats in the fetus is also important (Figure 4-3). Notably, non-DM mutation carrier siblings of CDM offspring have not been reported to suffer from any abnormal phenotypic features.

There is not yet an explanation for the exclusive maternal inheritance of CDM and it is not likely to be a simple one. The explanation may be found when the pathophysiology of the disease is better understood in terms of gene expression and metabolic pathways associated with it.

Parent of origin differences in the outcome of transmission of the DM mutation and implications for the mechanism of expansion of the CTG repeat

As seen in the previous chapter (Figures 3-2 and 3-5), parent-offspring mutation sizes show some degree of correlation and, as shown in Figures 4-1 and 4-2, DM mothers have, and have offspring with, a higher number of repeats in the mutant allele than DM fathers. In general, a higher number of CTG repeats and more severe clinical phenotypes are seen in offspring of maternal transmissions. However, in paternal transmissions intergenerational reductions in the number of CTG repeats are seen more often. More specific parental-gender related differences can be seen when the intergenerational mutation dynamics, as represented in Figure 3-1, are divided into paternal and maternal transmissions (Figures 4-1 and 4-2). *Paternal and maternal outcome of transmission of the same size mutations differ.* Reductions in the size of the mutation are more likely in paternal transmission of DM alleles containing over 700 CTG repeats.

There is little information concerning the stage(s) in intergenerational transmissions at which expansions and contractions occur. We were unable (either due through a

previous vasectomy or unwillingness of the relevant individuals) to analyse the size of the mutation in the sperm cells of the relevant DM fathers in order to see if the contractions were already present in those cells. Although both expansions and reductions in the number of CTG repeats seem to depend on the mutation size and gender of the transmitting parent they may not occur through the same mechanism(s). Parental rather than offspring dependent mutation dynamics may suggest that pre-zygotic components are relevant to the observed intergenerational changes in the DM mutation.

In conclusion, mutation size and the gender of the transmitting parent are seen to affect the intergenerational transmission dynamics of this mutation. Cell-type, mitotic, meiotic, and/or physiological differences between males and females may influence the dynamics of amplification and contraction of this CTG repeat.

Further studies will be needed in order to discern the possible role of each one of these components.

CHAPTER V

CHARACTERISTICS AND BEHAVIOUR OF THE MUTATION IN TISSUES FROM DM PATIENTS

1. Introduction

This chapter presents a summary of my analyses of the possible effects of mitosis and meiosis on the instability of CTG repeat mutation in DM. In order to study the possibility that the cell-cycle-linked DNA replication which occurs in mitosis affects the stability of the DM mutation, I conducted *in vivo* and *in vitro* analyses of various tissues from DM patients. Somatic and germ-cell, embryonic, new-born, and adult DM patient tissues, as well as tissues obtained at different stages within a lifespan, were studied in relation to mutation size changes.

Unlike individuals with "normal" chromosomes, DM patients showed heterogeneity in the number of CTG repeats in the mutant allele in the different tissues and within the same tissue. This heterogeneity seemed to derive from a mitosis-linked or a time-dependent tendency of the mutation to expand.

The possibility that meiosis in males could contribute to the intergenerational dynamics of the DM mutation was investigated. This was done by comparing the size of the

mutation in the sperm cells of DM fathers to the size of the mutation in peripheral blood leukocytes in their affected offspring.

Changes in the mutation were observed when cells obtained from affected tissues were kept in *in vitro* culture under conditions of exponential growth. These changes were seen as a steady increase in the size heterogeneity and in the number of CTG repeats.

The type of studies that could be conducted were generally limited, due to the difficulty in accessing key patient tissue materials. The conclusions derived from the studies that were carried out were, nevertheless, extremely useful in helping to identify some of the basic components which may be linked to the instability of this mutation.

2. Methodology

DNA analysis

The DNA obtained directly from the tissues of DM patients was extracted with an ABI automated DNA extractor according to the manufacturer's protocol. The DNA from cells in culture was extracted following the method of Birnboim (1992). PCR analysis and Southern blotting was conducted as described in the methods section in Chapter II. The probes used for band detection in the Southern blots were either pGB2.2 (in *EcoRI* restriction digests) or 750bpBamHI (in *BglI* restriction digests).

Tissue culture

Primary cell cultures originating from skin fibroblasts and amniotic fluid derived amniocytes were established. The tissues were obtained from human embryos with the DM mutation that were at their 3rd to 4th month of development. These adherent cells were maintained in an incubator at 37°C in the presence of 5% CO_2. The tissue culture medium consisted of alpha-MEM supplemented with 15% fetal calf serum as well as penicillin and streptomycin (all from GIBCO). Exponential cell growth was maintained through successive cell passages. Cell passages were performed by washing the cells in 1 x PBS (phosphate buffered saline) followed by exposure to trypsin (GIBCO) once they attained 90-95% confluence, and followed by replating at 1/4-5 split ratios.

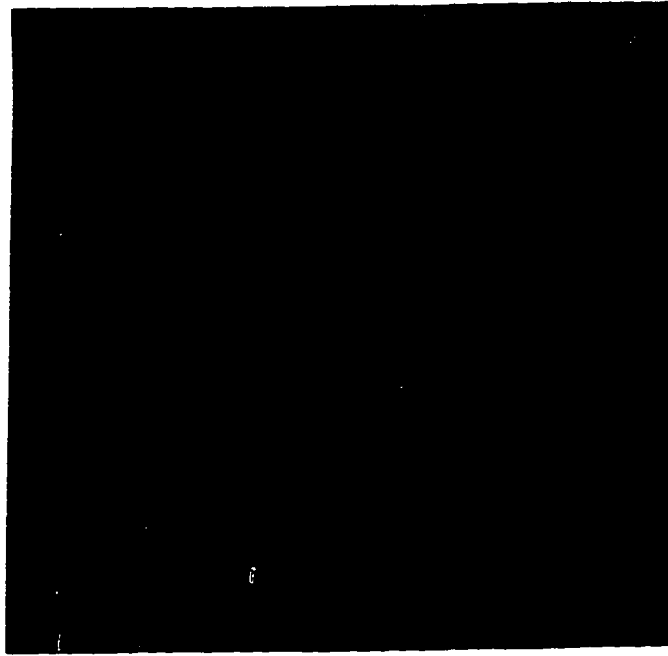
3. Results

Tissue distribution of mutant allele sizes

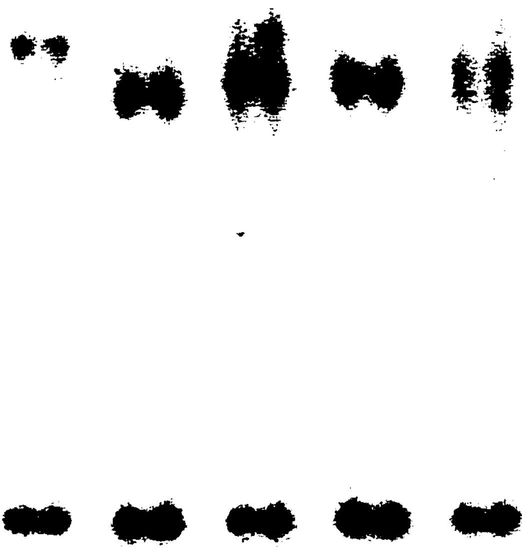
Analysis of mutation size was conducted on various tissues obtained from DM mutation carriers. The tissues originated from fetuses, infants, and adults with the disease. This analysis provided some indication of the dynamics of the mutation *in vivo* and of the possible relation between mitotic activity and instability of the CTG repeat. Figure 5-1 shows the results of some of these analyses. In general, and as shown, fetal tissues (A) from a fetus known to carry the mutation at the end of the first trimester in development present a less heterogeneous mutation size than newborn infants with the disease (B). In turn, adults with the disease have the greatest amount of mutant allele

Figure 5-1. Southern blot analysis of the status of the mutation in tissues from a fetus, a newborn infant, and an adult. **A** represents a fetus, **B** represents a new-born, and **C** represents an adult. The lower bands represent the normal allele and the upper bands represent the mutant allele. In **A** and **C** *EcoRI* restriction digestion preceded the Southern blot analysis whereas in **B** *BglI* digestion was performed instead. The different mobility and intensity of the bands shows the tissue variations in band size and heterogeneity. Letters on top of the lanes represent the tissue type. **A** represents tissues from a fifteen-week old fetus. **To** = tongue, **Te** = testicle, **S** = skin, **P** = pancreas, **M** = skeletal muscle, **I** = intestine, **L** = lung. **B** represents tissues from a new-born. **H** = heart, **B** = brain, **D** = diaphragm, **M** = skeletal muscle, **L** = liver. **C** represents spleen tissue from a 55-year-old DM patient. The lower arrow points to the normal allele and the upper arrow points to the heterogeneous and smeared mutant allele bands.

A



B



C

size heterogeneity. An example of this as can be seen in the spleen tissues from a 55-year old patient (C).

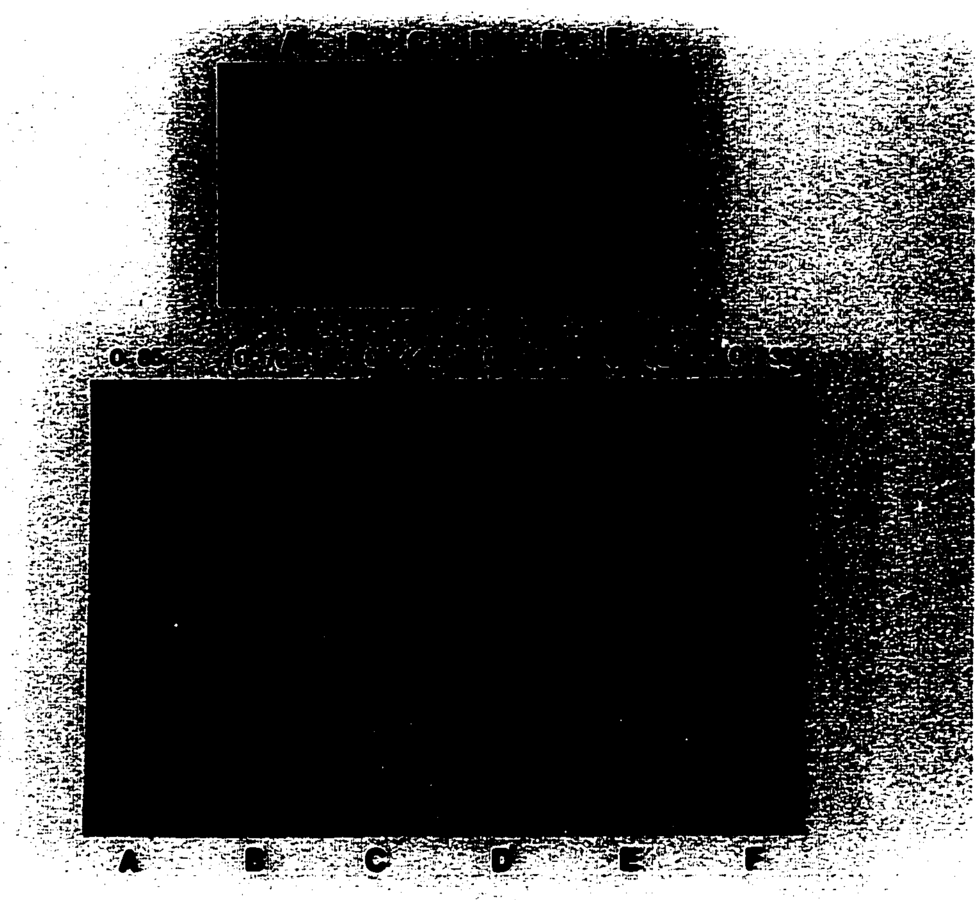
Mitosis or time-dependent variation in the size of the mutation

The assessment of the effect of age -- and therefore also possibly the involvement of mitotic activity -- on changes in the size of the DM mutation was performed by analysing the DNA of various individuals with DM. Given the technical limitations in the measurement of changes with the size of very small mutations, PCR analysis of relatively small size mutant alleles seemed the most appropriate technique to use in these experiments. Comparisons were made on the size of the mutation in peripheral blood leukocytes over a defined time period. Figure 5-2 shows the result of an experiment where the position of hybridization smears, which correspond to the mutant alleles, are compared. Molecular analysis was conducted on six DM patients with a number of CTG repeats in the mutant allele ranging from 60 to 400. Although not apparent in all samples, a slightly elevated smear on the right hand side lane in sets C, E and F, are suggestive of an increase in the number of CTG repeats during the defined time period.

Possible effect of meiosis on mutation size changes: size distribution of the mutation in sperm cells

A basic event involved in the intergenerational transmission of genetic material is meiosis. In order to assess the possible role of male meiosis in the intergenerational

Figure 5-2. Expansion of the DM mutation in patient's leukocytes over a time period. This figure shows the results of the PCR analysis of possible increases in the number of CTG repeats in small-size mutations. The top panel shows the samples in an agarose gel stained with ethidium bromide, the bottom panel shows the same products after Southern blot transfer and probing with a $(CAG)_{10}$ oligonucleotide probe. **A** to **F** represent six different individuals. The number of CTG repeats in their mutant alleles ranges from approximately 60 to 400. The lower bands correspond to the normal alleles (not seen in **A**, due to inefficient hybridization of the probe). The smeary appearance of the bands indicates the inefficiency of a standard PCR reaction on mutant DM alleles. The left lanes in **A** to **F** give an indication of the number of CTG repeats in the mutation by the relative position of the smear. The right lane is equivalent except that the DNA used as template in the PCR reaction was obtained from the patient a number of months later, as indicated by the number on the top of the respective lanes. The slightly upward shift of the smear in some of the right-hand side lanes suggests somatic instability.



expansions and contractions of the mutation, a comparative tissue study of the size of the mutant CTG repeat from four DM individuals who presented approximately 80-100 CTG repeats in their blood leukocytes was conducted. The size of the mutation in these tissues was compared to the size of the mutation in the leukocytes of their DM offspring. In all cases, sperm cells showed an average increased size and a very high heterogeneity of the mutant allele compared to the same allele in peripheral blood leukocytes. This heterogeneity was seen as a smear in both PCR and genomic DNA Southern blots. Figures 5-3 and 5-4 show a representative example of one of the kindreds analysed for this purpose. Although the bands are faint, the genomic Southern blot (derived from a *BglI* restriction digest) is also included to provide a better appreciation of the allele sizes present in genomic DNA (Figure 5-4). *Sperm cells in these individuals showed a much increased size and heterogeneity of the mutant allele than blood leukocytes.* The size of the mutant allele in leukocytes in their progeny showed great variations, including offspring with an average mutant allele size possibly smaller than the average mutation size in the sperm of the father (Figure 5-3, lane 6).

Possible effect of mitosis on DM mutation size changes: tissue culture study

In order to assess the possibility that this mutation is not stable in mitotically active cells, tissue culture was conducted on skin fibroblasts from a newborn infant with CDM and amniocytes obtained from a DM fetus. Serial passages were performed in order to allow for exponential growth. A subset of these cells was selected at each passage and used for DNA analysis. The results of one of these experiments is shown in Figure 5-5.

Figure 5-3. Comparative PCR analysis of mutant allele sizes in an individual's blood leukocytes, sperm cells, and blood leukocytes of his offspring. Products of PCR amplification of the mutation on an agarose gel (**A**) and following Southern blotting and hybridization with a $(CAG)_{10}$ oligonucleotide probe (**B**). This pedigree represents the transmission of a mutant allele from an affected father (black square top left) to four of his five offspring. **g** refers to the germ-line (sperm) cells and **s** refers to the somatic cells as represented by blood leukocytes. The PCR reaction was performed with mutation flanking primers 406 and 409. An alkaline phosphatase linked $(CAG)_{10}$ oligonucleotide was used as a probe in the Southern blot.

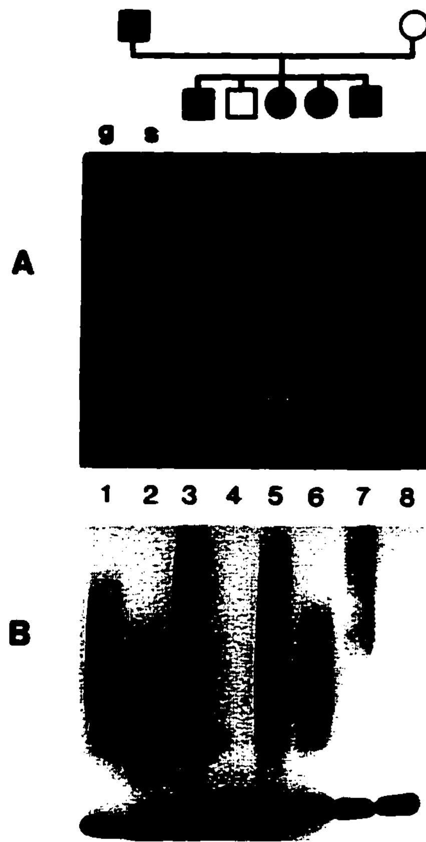

g
s
A
1
2
3
4
5
6
7
8
B

Figure 5-4. Comparative genomic Southern blot analysis of mutant allele sizes in an individual's blood leukocytes, sperm cells, and blood leukocytes in his offspring. (Same pedigree as in Figure 5-3). Autoradiograph of a genomic DNA Southern blot corresponding to a *BglI* restriction digest probed with the probe 750bpBamHI. Pedigree representing the transmission of a mutant allele from an affected father (black square top left) to four of his five descendants. **g** refers to the germ-line (sperm) cells and s refers to the somatic cells as represented by blood leukocytes. The mutant alleles are shown as the upper bands or smears in lanes **1-3** and **5-7** whereas the normal alleles are the lower compact bands.

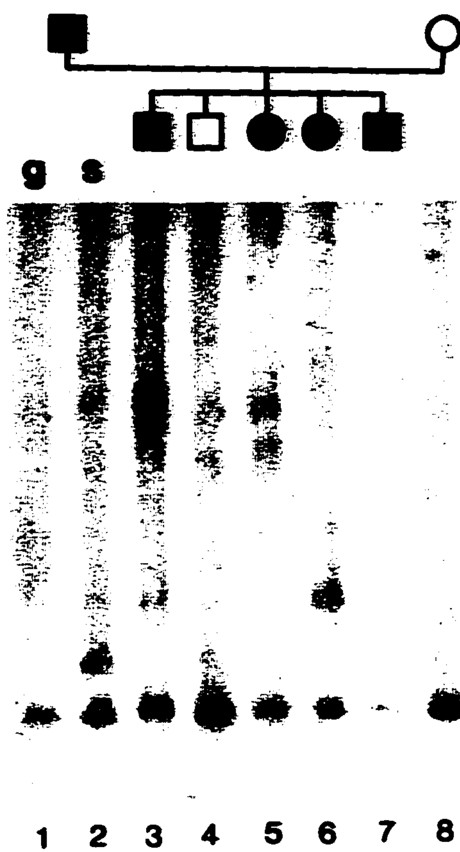
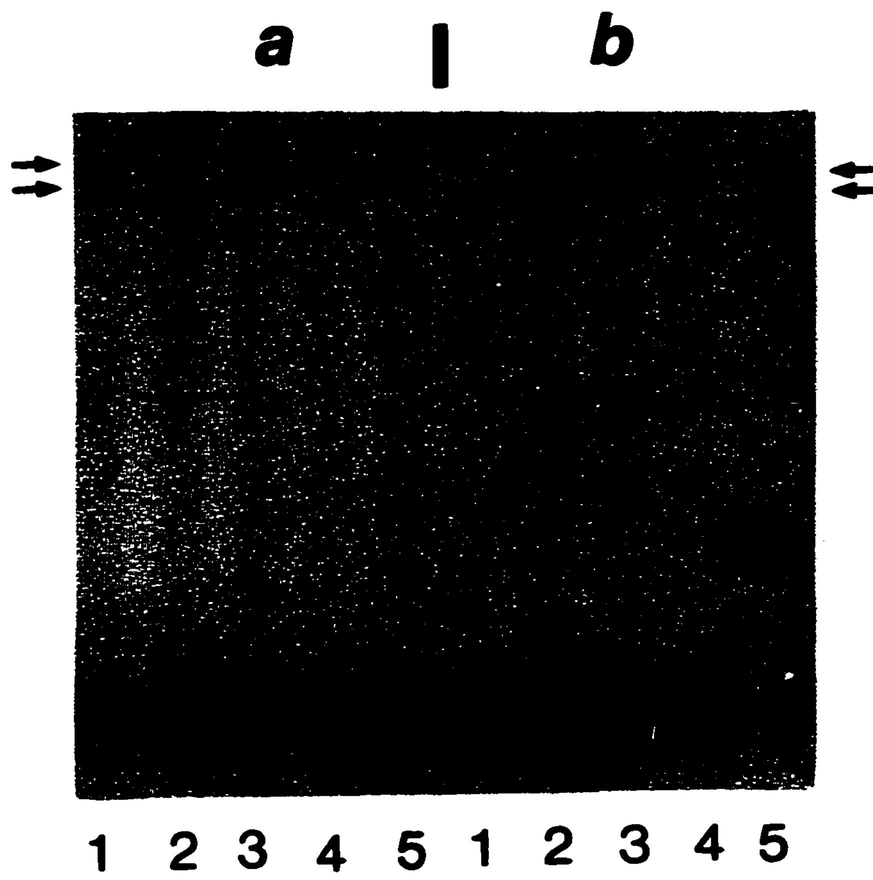

g
s
1
2
3
4
5
6
7
8

Figure 5-5. Study of the behaviour of the mutant allele in tissue culture. *BglI* Southern blot representing the tissue cultured amniocytes (**a**), and fibroblasts (**b**), from a fetus and a new-born respectively. Arrows point to the bands corresponding to the mutant allele (of approximately 1800 CTG repeats in this case). Lower bands correspond to the normal alleles. Lanes **1-5** represent the passages following primary cell culture. The mutation does not appear to be stable through the mitotic division of these cells in culture. The changes seem to involve a slight increase in size and heterogeneity of the mutant allele band.


a
b
1 2 3 4 5 1 2 3 4 5

BglI restriction digestion of the DNA from those cells at each passage, followed by gel electrophoresis, Southern blotting and probing with the 750bpBamHI probe showed a slight increase in the size of the mutant allele (upper band). This increase seemed to be coupled with a decrease in intensity and an increase in heterogeneity of the same allele band.

4. Discussion

Studies such as those described in this chapter are limited by the fact that key patient material is difficult or practically impossible to access. This limits the conclusions that can be drawn in relation to the *in vivo* effects of mitotic and meiosis. The results presented here, nevertheless, were confirmed by others with respect to somatic heterogeneity of the mutation (Anvret et al., 1993; Thorton et al., 1994a; Zatz et al., 1995), the increase in size and heterogeneity of the mutation with number of CTG repeats and with age (Wong et al., 1995), the possible meiotic instability (Brunner et al., 1993a; Monckton et al., 1995), and the *mitotic instability of the mutant CTG repeat in cells grown under tissue culture conditions* (Hecht et al., 1993; Abeliovich et al., 1993; Wöhrle et al., 1995).

A tendency for this mutation to expand in somatic tissues from patients was observed. A direct example of this was obtained through the observation of slight mutation size increases, *in vivo*, in peripheral blood leukocytes with time, in some patients (Figure 5-

2). At the time when these tests were done there were technical limitations associated with an accurate assessment of small changes in the size of DM mutant alleles. Accessibility of appropriate DNA samples from patients was limited and therefore comparisons became difficult; nevertheless, some indication of subtle but definite increases in the size of the mutant allele were obtained and later confirmed independently by Wong et al., (1995), and others. These tests did not distinguish between expansion due to possible slippage during the cell-cycle-linked DNA replication of the CTG repeat sequence in mitosis and expansion due to other processes.

The fainter and more heterogeneous mutant allele bands and smears corresponded to DNA extracted from older DM patients. Direct analyses of tissues suggested increased mutation size heterogeneity with age (see Figure 5-1 for an example). This result is in contrast with what has been observed in the mutation responsible for the majority of fragile X syndrome cases (Rousseau et al., 1991). These results suggest an *increase in size and heterogeneity of the mutant allele band that occurs as a mitosis-linked or as a time-dependent process in cells from DM patients.*

The size of the mutation was also seen to vary with tissue type, as shown in Figure 5-1. Nevertheless, the variety of tissues and the number of different tissues from DM patients was insufficient to assess the consistency of the size heterogeneity patterns. If some degree of consistency in relation to the dynamics of the mutation in the different tissues was present, it would be possible to compare or seek a correlation between

mitotic activity (and therefore cell-cycle-linked DNA replication) and expansion of this mutation *in vivo*. Interestingly, nevertheless, it was observed that mutant allele band size heterogeneity was increased in older DM patients (in DNA obtained from leukocytes). Moreover, the comparison of the size of the mutation in skeletal muscle and leukocytes in adult DM patients (Anvret et al., 1993; Thornton et al., 1994a; Zatz et al., 1995) showed considerably larger expansions in muscle cells. This is suggestive (together with Figures 5-1 and 5-2) of a mechanism of expansion that is not necessarily linked to the mitotic activity of the cell.

The observed allelic instability in somatic tissues would be expected to interfere with a study of the possible effect of meiosis on DM mutation dynamics on tissues obtained from patients. I proceeded, nevertheless, to analyse the sizes of the mutation in sperm cells and peripheral blood leukocytes of DM patients and to compare them to the sizes of the mutation in their offspring. DM males with 80-100 CTG repeats in the mutant allele in blood leukocytes had larger and more heterogeneous mutant alleles in sperm cells. Their DM offspring also showed large CTG repeat expansions. Mutant alleles in the sperm cells of parents were reported to be similar in size to those found in peripheral blood leukocytes in their offspring (Brunner et al., 1993a). In the cases that I studied, it appears that an important part of the intergenerational expansion of the mutation observed in peripheral blood leukocytes could already be seen in the terminally differentiated germ-cells of the fathers. A representative example of our results is shown in Figures 5-3 and 5-4 where the average size of the mutation in

peripheral blood leukocytes in DM offspring with the mutation is comparable to the average size of the mutation in the father's sperm cells. Since the differences in the size of the mutant alleles in siblings are often quite pronounced, it is possible that the intergenerational expansions depend, considerably, on the size of the mutant allele passed on by the affected parent. Moreover, the fact that the bands representing the mutant allele show relatively low heterogeneity within and among fetal tissues (Figure 5-1 (A)), also favours the argument of *pre-zygotic intergenerational changes being important components in the intergenerational dynamics of this mutation.*

Our results are in contrast with those found in fragile X, where the full mutation in the FMR1 gene in affected males is absent in their sperm (Reyniers et al., 1993). In this case post-zygotic events may be more likely to underlie the intergenerational dynamics of the mutation. Unlike in DM, mutation size heterogeneity within tissues in individuals with the fragile X syndrome is often manifested as a few major discrete bands that can, in some cases, already be seen in fetal tissues (Devys et al., 1992). Moreover, in DM, and in contrast to equivalent studies in the fragile X mutation (Wöhrle et al., 1993) and in the CAG repeat responsible for SBMA (Table 1-1), (Spiegel et al., 1996), fetal tissues in culture show clear changes in the size of the mutant allele. These observations (together with studies of monozygotic twins) may suggest that, perhaps unlike in DM, the major events involved in the observed intergenerational expansions of the CGG repeat in the fragile X syndrome are due to mutational processes occurring at a very early stage in development. Since the DM and the fragile X mutations share many

features (see Table 1-1), if there is a common model for the mechanism of expansion for these mutations it must account for those differences.

Meiosis leading to the formation of mature germ cells involves at least 10 times more mitotic cell divisions in males than in females. Since mothers do not pass on smaller mutant DM alleles than fathers (Chapter IV) it is possible that mitosis itself is not a main determinant in the expansion of the DM mutation. It may, nevertheless, perhaps be linked to the higher percentage of intergenerational reductions in the number of CTG repeats seen in paternal transmissions.

In conclusion, these studies suggest that DM mutation's instability is seen in cells that undergo mitosis and meiosis, although cell-type and other unknown elements may affect more directly the dynamics of the mutation. This conclusion is based on the fact that although there is a tendency for this mutation to expand, the rate of expansion does not necessarily correlate with the rate of mitotic activity of the tissue (Anvret et al., 1993; Lavedan et al., 1993a; Thornton et al., 1994b; Zatz et al., 1995; and Figure 5-1).

The differences in the size of the mutation in sperm cells as compared to somatic cells in the progeny of DM fathers seems to indicate that pre-zygotic as well as post-zygotic components may be important determinants of the observed intergenerational expansions, and especially contractions.

The observed expansion of the DM mutation in tissue cultured cells (at least in fibroblasts and amniocytes), (Figure 5-5), provides a system in which to study triggers of instability. In such a system it is possible to test the possible link between cell-cycle-coupled DNA replication and instability of this mutation. This is studied in the next chapter.

CHAPTER VI

THE CELL-CYCLE-LINKED, DNA-REPLICATION-INDEPENDENT MECHANISM OF EXPANSION OF THE DM MUTATION: A STUDY OF THE POSSIBILITY OF DNA REPAIR AS THE BASIS FOR INSTABILITY

1. Introduction

Although the mechanisms are unknown, models have been proposed which may explain the highly unstable nature of the trinucleotide repeat sequences that cause human disease (Richards and Sutherland, 1992, 1994; Caskey et al., 1992). These models are based on slippage of either DNA polymerases, Okazaki fragments, or the leading or lagging strands of the repeat sequence itself during DNA replication. Microsatellite instability was observed in the tumours of individuals carrying mutations in DNA mismatch repair genes (Fishel et al., 1993; Leach et al., 1993; Bronner et al., 1994; Papadopoulos et al., 1994). However, to date, there is no evidence of the involvement of these or any other mechanisms in the high instability of the known microsatellites that cause human disease (Table 1-1).

The observation of DM mutation instability in cells from patients in culture (Chapter V) enabled us to examine the possible involvement of cell-cycle-coupled DNA replication in this process. I conducted *in vitro* tissue culture experiments with cells

obtained from DM patients. Analysis of possible chromosomal DNA replication-dependent expansion of the mutation was performed by a parallel comparison of mutant CTG repeat array length in mitotically active and quiescent fibroblasts and myoblasts. The cells originated from primary cultures of fetal tissues taken at different developmental stages and containing 700 or 1800 CTG repeats in the mutant allele.

Previous reports had shown expansion of this mutation when cells derived from DM embryos were grown in these conditions (Hecht et al., 1993; Abeliovich et al., 1993; Wöhrle et al., 1995). Interestingly, I also observed the same phenomenon when contact inhibition impaired the capacity of these adherent cells to proliferate. Such results questioned the role of chromosomal DNA replication as the primary cause of expansion of this mutation. In the proposed slippage models DNA replication is key, yet in contact inhibited cells the only replication expected to take place is the DNA replication associated with DNA repair. Following this observation I focused on events that could explain the expansion of this mutation in the absence of cell-cycle-coupled DNA replication.

I focused on DNA repair as a possible cause of instability. Since DNA breaks are intermediates of DNA repair processes, assays for the presence of DNA breaks and gaps within the mutation were performed. For this purpose, cell culture was conducted in the presence of DNA repair inhibitors. The possible presence of breaks was studied through semi-quantitative PCR measurement of blocks to the polymerase in the reaction.

Analysis of break type (whether single- or double-strand break) was done by S1 nuclease sensitivity analysis, and denaturing and non-denaturing gel electrophoresis.

2. Methodology

Cell culture

Primary cultures of cells originating from fetal tissues were established and maintained in an incubator at 37°C and 5 % CO_2. The tissue culture medium consisted of D-MEM supplemented with 15% fetal calf serum and antibiotics for fibroblasts or F12 supplemented with 20% fetal calf serum and antibiotics for myoblasts (all from GIBCO). The 15% fetal calf serum was substituted by 3% calf serum (GIBCO) for contact inhibited cells which were later treated with DNA repair inhibitors, and in long-term confluent cultures. Exponential cell growth was maintained through cell passaging at split ratios of 1/4-5. In each passage cells were allowed to grow to 90-95% confluence. Viability of cells that had been maintained in long-term confluent cultures was determined through FACS (fluorescent activated cell sorting) and proliferation curve analysis following trypsinization and cell replating.

Chemicals

Stock solutions of cytosine arabinoside (Ara C), adenine arabinoside (Ara A), and aphidicolin (all from Sigma) were prepared by dissolving these chemicals to appropriate concentrations in 1 x PBS.

DNA analysis

Genomic DNA from the cells in culture was extracted with the use of a phenol-chloroform extraction protocol according to the instructions provided with the Applied Biosystems 341 Nucleic Acid Purification System. The oligonucleotide pair DMK9003 (5'-CACAGGCTGAAGTGGCAGTTCCA-3') and DMK11111 (5' - TGTCGGGGTCTCAGTGCATCCA-3'), (DMK gene, GeneBank accession number L00727) bracket 2.1 kb of DMK gene sequence around the mutation. They were used as primers in PCR reactions as described (Cheng et al., 1996), in 50 μl reactions that consisted of 1 x GeneAmp® XL PCR Buffer (Tricine, K-acetate, glycerol, and DMSO); 0.2 mM each of dATP, dCTP, and dTTP; 0.2 mM total of dGTP and 7-deaza-dGTP; 0.2 μM of each primer; 1 U of rTth DNA polymerase, XL; 1.2 mM $Mg(OAc)_2$; and approximately 15-50 ng of total genomic DNA. All PCR reagents were from Perkin-Elmer, except for the 7-deaza-dGTP (Boehringer Mannheim). The percentage of 7-deaza-dGTP in the dG/deaza-dGTP reaction component varied from 25 to 50 % depending upon the size of the repeat expansion. Conditions approaching quantitative detection of PCR products were achieved by serial collection of samples in the late exponential phase of the reaction. Amplifications were performed in a GeneAmp® PCR thermal cycler System 9600 (Perkin-Elmer). The initial template denaturation step was 15 sec at 95°C followed by 22-30 cycles of denaturation at 95°C for 10-12 sec, followed by annealing and extension at 67°C for 6 min. A final incubation at 68°C for 10 min was used for completion of strand synthesis. The products were loaded in 0.8% agarose gels without ethidium bromide and submitted to electrophoresis for 3 h (DNA

break analyses) or 30 h (mutation size analyses). The gel analysis and Southern blot detection of products was performed as described in Chapter II. Quantitative measurement of signals on autoradiographs was performed with a Bio-Rad Model GS-670 Imaging Densitometer. The possible presence of widespread breaks in DNA treated with the repair inhibitors was analysed through gel electrophoretic analysis of the mobility of whole genomic DNA and quantitative long- and short-range PCR on various sequences at 19q13.3 and 7q31.

S1 nuclease sensitivity assay and denaturing gel electrophoresis

A series of 10 μg aliquots of DNA were digested with 50 U of S1 nuclease (GIBCO) in 40 mM potassium acetate (pH 4.6), 300 mM NaCl, and 1.3 mM $ZnSO_4$ for 2 to 10 min at 45°C. The reaction was terminated by adding EDTA to a final concentration of 5 mM and followed by heat inactivation at 70°C for 10 min. The DNA was then extracted with phenol-chloroform and ethanol precipitated. This DNA was then digested with *EcoRI*, submitted to denaturing (NaOH) agarose gel electrophoresis, Southern blotted, and probed with pGB2.2 as described.

FACS and [^{14}C]-thymidine incorporation analyses

FACS analysis was preceded by cell permeabilization with the use of a standard protocol involving ethanol and acetic acid treatment followed by propidium iodid staining. FACS analysis was performed through the use of a fluorescent activated cell sorter from Coulter electronics and the data were analysed with XL® and Multicycle®

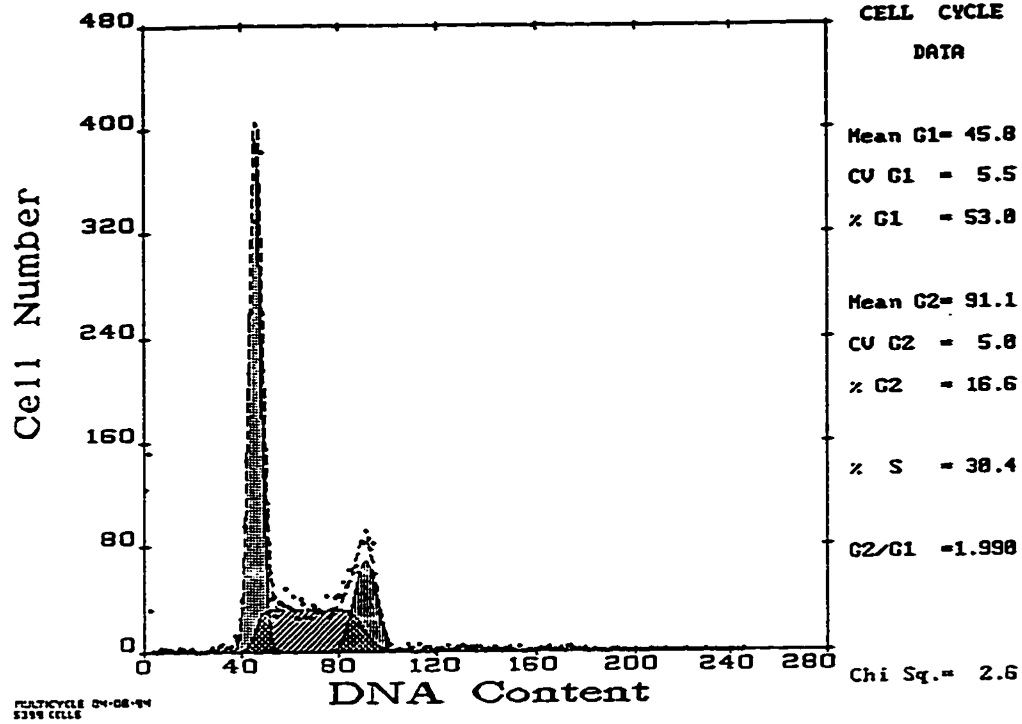
(Phoenix Flow software products) software on a standard PC. Assessment of cell cycle status was done on sets of approximately 5 x 10^4 actively cycling and confluent cells. A complementary DNA replication status test was performed by measuring the incorporation of [^{14}C]-methyl thymidine (Amersham) into control cycling and confluent cells in culture. The release of DNA for this test was performed by incubating the cells at 37°C overnight in 0.2 N NaOH, 0.5% SDS, and carrier DNA. This was followed by the addition of HCl and incubation at 0°C. Radioactivity was measured following resuspension of the pellet in NH_4OH and readings with a Beckman LS-7000 scintillation counter.

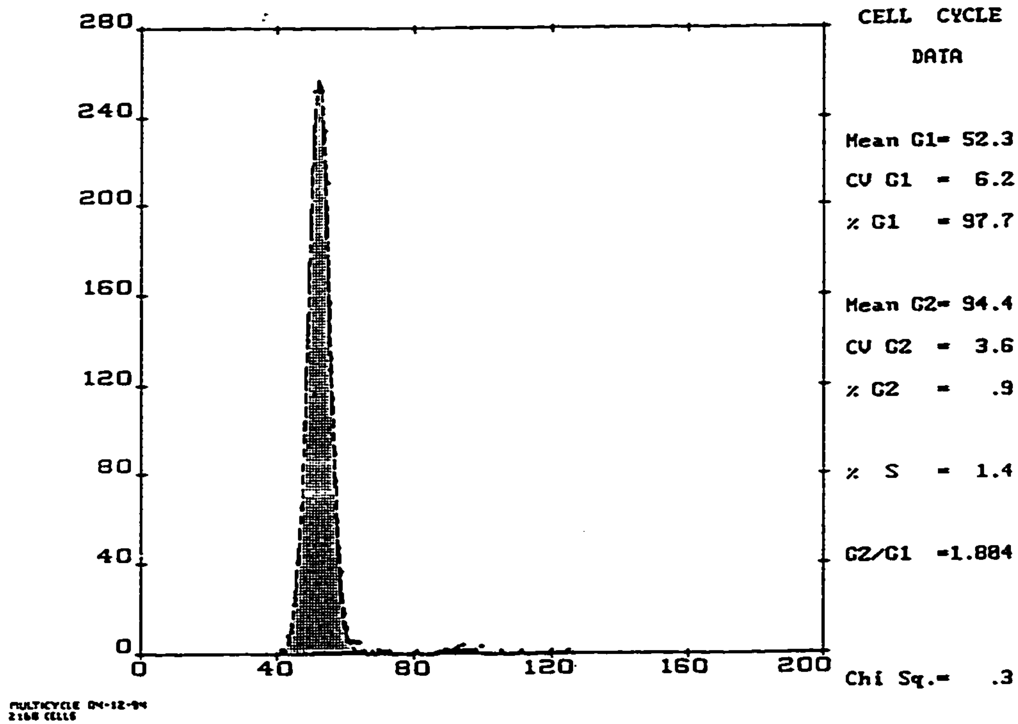
3. Results

Cell cycle analysis

Cycling and 100% confluent (exhibiting the typical swirling pattern) cells in culture were trypsinized and immediately permeabilized, stained, and submitted to FACS analysis. FACS analysis in confluent cells revealed one only sharp peak corresponding to cells at G0 within the cell cycle (>99% of the population), (Figure 6-1 bottom as an example). The simultaneous FACS analysis of the same cells, although non-confluent or cycling, maintained in parallel tissue culture conditions showed a regular cell cycle distribution (Figure 6-1 top, as an example).

Figure 6-1. Cell cycle profile obtained through FACS analysis of non-synchronous cycling and confluent (synchronous at G0) fibroblasts. Top figure shows the distribution of a population of asynchronous cells through the different stages of the cell cycle. The bottom figure shows the profile of synchronous cells at G1 at 97.7 % of the (also referred to as G0 when the cells are at this stage because of cell cycle withdrawal).


480
400
320
240
160
80
0
Cell Number
0
40
80
120
160
200
240
280
DNA Content
CELL CYCLE
DATA
Mean G1= 45.8
CV G1 = 5.5
% G1 = 53.8
Mean G2= 91.1
CV G2 = 5.8
% G2 = 16.6
% S = 38.4
G2/G1 =1.998
Chi Sq.= 2.6

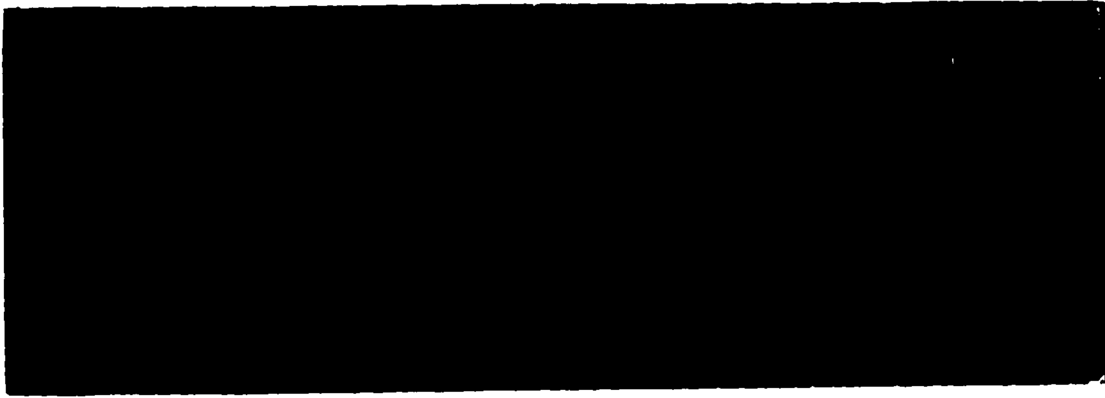

280
240
200
160
120
80
40
0
0
40
80
120
160
200
CELL CYCLE
DATA
Mean G1= 52.3
CV G1 = 6.2
% G1 = 97.7
Mean G2= 94.4
CV G2 = 3.6
% G2 = .9
% S = 1.4
G2/G1 =1.884
Chi Sq.= .3

A cell-cycle-linked DNA-replication-independent mechanism of expansion of the DM mutation

The study of a possible link between DNA replication and amplification of the CTG repeat was undertaken by comparing changes in the mutation in mitotically inhibited fibroblasts and myoblasts from DM individuals to the same cells after a period of normal proliferation. Early passage cell cultures from fetal tissues were established, brought to confluence, and divided into three sets. The first set was used for immediate DNA analysis. The second contained cells that at this point were replated at a lower density and allowed to proliferate at an exponential rate. The third consisted of the original contact-inhibited cells which were not allowed to proliferate because of contact inhibition during the same time-period in which the other set of cells were proliferating. Verification of cell proliferation and cell cycle status was performed through [^{14}C]-thymidine incorporation into DNA and flow cytometric (FACS) analysis. The contact-inhibited cells remained as a monolayer and were found to be in G0 (>99%) as determined by FACS, and incorporated virtually no [^{14}C]-thymidine compared to those that were actively proliferating.

The results of this experiment are shown in Figure 6-2. The bands represent the mutant allele (of approximately 700 CTG repeats) in these cells, and were generated by PCR using the described DM mutation flanking primers. The products were submitted to long agarose gel electrophoresis in order to increase the resolution of the mutant bands. The normal allele band (2 kb) had, at this point, run off the gel. The bands are of 4 kb

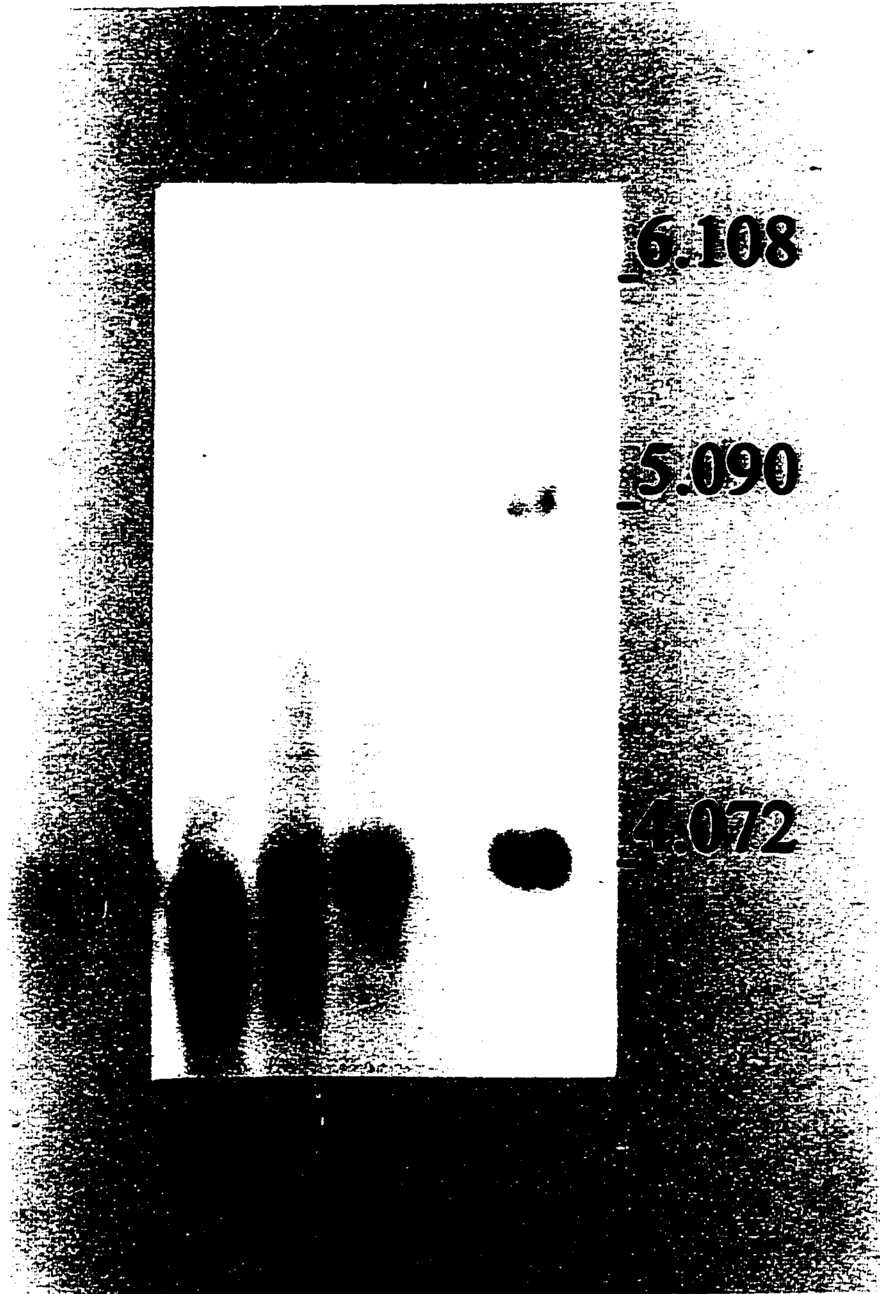
Figure 6-2. Analysis of expansion of the mutation through active cell cycling and through contact inhibition. DNA sequences produced by PCR following gel electrophoresis, Southern blotting membrane transfer, and probing with a $[^{32}P]$-$(CAG)_{10}$ oligonucleotide are shown. The template DNA originated from DM cells in which the number of CTG repeats in the mutant allele is approximately 700. The products were submitted to long agarose gel electrophoresis in order to increase the resolution of the mutant band. The band in this figure contains 4 kb of sequence of which 2.1 kb correspond to the CTG repeat. Lanes **1** and **7** represent the size of the mutation following cell growth to contact inhibition (control). Lanes **2, 3, 4** and **6** represent four consecutive cell passages following contact inhibition, lane **5** represents cells held as confluent cultures (contact inhibited) during the time period needed for four successive cell passages. The mobility of the bands indicates an increase in size of the mutation in the long-term contact inhibited cells (lane **5**) compared to the same cells at the start of this period (lanes **1** and **7**).


1 2 3 4 5 6 7

of which 2.1 kb correspond to the mutation sequence. The mobility of the bands indicates an increase in size of the mutation in contact inhibited cells (lane 5) compared to the same cells at the start of this period (lanes 1 and 7). The size of the band in passage 4 cells indicates that the acquisition of CTG repeats was slightly higher in the cells that were mitotically active during the same time period. The actively replicating cells in culture, at approximately 2.5 cell divisions per passage, showed an increase in the number of CTG repeats relative to the number present at the start of the culture (lane 1) as had already been reported, and was within observed margins (Wöhrle et al., 1995). This increase (although slightly smaller) was also apparent in cells that were held in culture for the same time period but were mitotically quiescent (lane 5).

For a better appreciation of band shifts indicative of changes in the number of repeats in the absence of cell-cycle-linked DNA replication, cells in culture were maintained at confluence for a longer time period. Figure 6-3 shows the result of an experiment in which contact inhibition was allowed to proceed for approximately two months. The bands contain the mutation sequence flanked by 2.1 kb of the DMK gene. One can clearly see the increase in the number of repeats that occurred during this time period when the mutant allele band size was compared between cells at the beginning and at the end of this period at confluence. In the cells represented in figure 6-3 this increase during 65 days at confluence is approximately 30-50 CTG repeats.

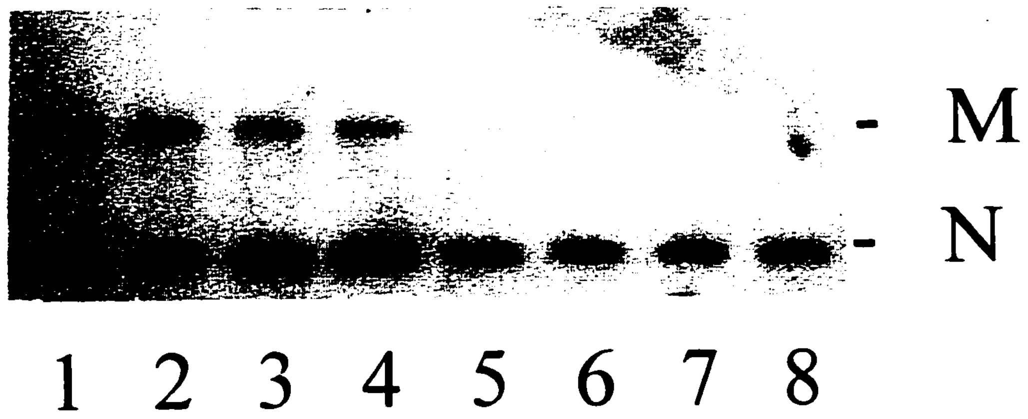
Figure 6-3. Analysis of the extent of expansion acquired by cycling and contact-inhibited cells. The bands are equivalent to those shown in the previous figure (same cells and size). In this case the cells had been contact-inhibited for the number of days indicated on top of each lane. Lane **1** corresponds to the size of the mutation immediately following cell growth to contact inhibition. Lanes **2** and **3** show the bands obtained after the same cells had been kept (further) contact-inhibited for the number of days indicated on top of their respective lanes. The ladder in the most right had side lane indicates the size of the bands.


6.108
5.090
4.072

Detection of DNA repair intermediates within the mutation

The expansion of the mutation which we observed in DNA replication-inhibited, or quiescent, cells led us to consider the possibility that DNA repair was involved in this process. The same, confluent, cells were kept in culture in the presence of different amounts and combinations of aphidicolin (as a DNA polymerase α and δ inhibitor) and the DNA repair inhibitors cytosine arabinoside (Ara C) and adenine arabinoside (Ara A) (as chain terminators). Figure 6-4 shows the presence of DNA repair intermediates, such as gaps or breaks, within the mutation. The breaks were detected with a semi-quantitative PCR assay. This highly sensitive technique for DNA break and gap detection (Grimaldi et al., 1994; Pogribny et al., 1995) is based on the ability of a strand-break to halt the progression of DNA polymerase. Cells were cultured for 7 days in the presence of different amounts of the described inhibitors (lanes 2 to 8). Following cell harvest, analysis of the mutation was performed. A shorter interval of electrophoresis prevented excessive separation of the normal (N, 2.1 kb) and mutant (M, 4.2 kb) allele bands. N was used as an internal control. Increasing the amounts of the described repair inhibitors resulted in gradual reduction of mutant allele band intensity. This is not paralleled by the reduction of the intensity of the normal allele band.

Figure 6-4. Detection of breaks in the DNA of cells treated with DNA repair inhibitors. Semi-quantitative PCR analysis of the mutation in untreated cells and cells treated with DNA repair inhibitors. Band **M** represents the same band as the one seen in Figure 6-3. Shorter time-interval gel electrophoresis resulted in the close migration of the corresponding normal allele band (N), which contains the same sequence as **M** except with a normal (<35) instead of a mutant (approximately 700 in this case) number of CTG repeats. All cells were brought to confluence previous to treatment with the inhibitors. Lane **1** represents the bands of an untreated subset of cells and was used as a control. Lanes **2** to **8** represent the bands of the same cells following exposure for 7 days to the following: lane **2**: 2 μM aphidicolin; lane **3**: 2 μM Ara A; lane **4**: 2 μM Ara C; lane **5**: 2 μM aphidicolin + 2 μM Ara A + 2 μM Ara C; lane **6**: 20 μM Ara C; lane 7: 100 μM Ara C; lane **8**: 5 μM Ara C + 2 μM Ara A + 2 μM aphidicolin. The disappearance of the bands at increasing inhibitor concentrations is not parallelled by the disappearance of the normal allele band which served in this case as an internal control. Above a certain concentration of repair inhibitor, polymerization of the mutant allele band is impaired.

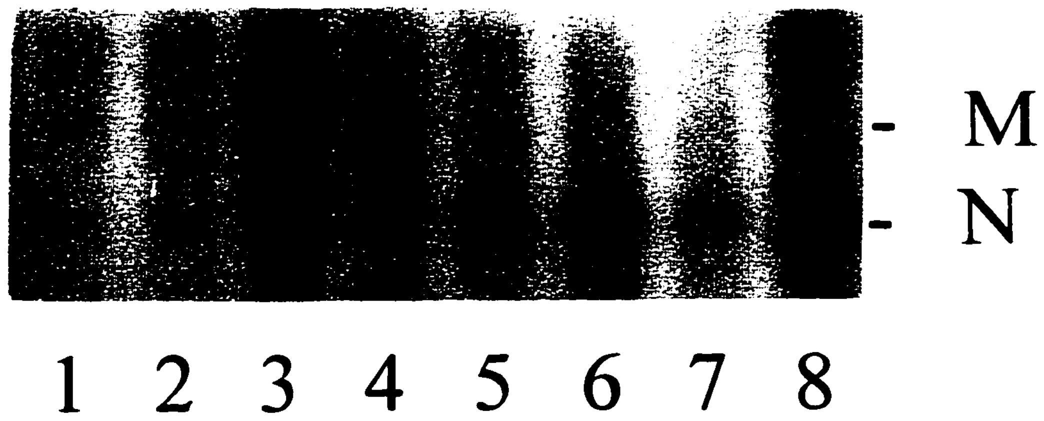

M
N
1 2 3 4 5 6 7 8

Disappearance of bands with this technique reveals the presence of obstacles to chain elongation such as breaks or gaps. These obstacles must be within the mutation since this is the only difference between bands M and N. In order to rule out an alterative explanation for the observed results, the integrity of the total genomic DNA from the same cells cultured for the same time period under a relatively high concentration of Ara C (1 mM) was studied. Gel electrophoretic analysis of total genomic DNA and quantitative long- and short-range PCR on sequences close to (yet excluding) the DM CTG repeat and in other various regions of the genome (at 7q31) were performed and revealed no significant differences compared to the control DNA (template DNA in Figure 6-4, lane 1) (data not shown).

Differentiation between single-stranded breaks (SSBs) and double-stranded breaks (DSBs)

In order to distinguish between DSB and SSB, *EcoRI* restriction digestion of whole genomic DNA, obtained from cells treated with DNA repair inhibitors, was performed and followed by gel electrophoresis, Southern blotting, and probing with pGB2.2. Figure 6-5 shows the results of this analysis. N and M represent the normal (approximately 9 kb) and mutant allele (approximately 12 kb) bands respectively. They were obtained from the same genomic DNA used as template in the PCR-based break detection assay (Figure 6-4). There is a gradual, albeit slower, reduction in the intensity of the mutant allele band at increasing concentrations of inhibitors. This indicated the possibility that DSBs, closely spaced SSBs, or overlapping gaps on complementary

Figure 6-5. Detection of double-stranded breaks (DSBs) in the DNA of repair-inhibitor treated cells. Genomic DNA Southern blot analysis of the mutation in untreated cells and cells treated with DNA repair inhibitors. Band **M** (12 kb) represents the DM mutant allele containing approximately 700 CTG repeats and flanked by approximately 10 kb of DMK sequence (including a 1 kb insertion polymorphism) bracketed by *EcoRI* restriction sites. Band **N** (of approximately 9 kb) is the corresponding band for the normal allele. Probing was performed with pGB2.2. All cells were brought to confluence prior to treatment with the inhibitors. Lane **1** represents the bands of an untreated subset of cells and was used as a control. Lanes **2** to **8** represent the bands of the same cells following exposure for 7 days to the following: lane **2**: 2 μM aphidicolin; lane **3**: 2 μM Ara A; lane **4**: 2 μM Ara C; lane **5**: 2 μM aphidicolin + 2 μM Ara A + 2 μM Ara C; lane **6**: 20 μM Ara C; lane **7**: 100 μM Ara C; lane **8**: 5 μM Ara C + 2 μM Ara A + 2 μM aphidicolin. The disappearance of the bands at increasing inhibitor concentrations is not parallelled by the disappearance of the normal allele band which served in this case as an internal control. The rate of disappearance of the mutant allele band is slower than in Figure 6-4 and therefore possibly indicative of the presence of SSBs as well as DSBs.


M
N
1 2 3 4 5 6 7 8

strands within the mutation were present. In order to further distinguish between these possibilities, denaturing gel electrophoresis analysis was performed. The fainter appearance of the mutant allele band in the DNA originating from inhibitor-treated cells indicated that the majority of breaks were on one of the two strands. S1 nuclease treatment of the same products followed by non-denaturing gel electrophoresis led to a similar result (data not shown).

4. Discussion

Finding DNA breaks

The results presented confirm the observed *instability of the DM mutation in cells from DM patients in culture and showed that this instability was not dependent on the cell-cycle-linked DNA replication* but was time-dependent, at least in cells from patients cultured *in vitro*. The use of a PCR technique that is efficient in the amplification of the DM long CTG repeat arrays (Cheng et al., 1996) greatly facilitated the analyses. The use of DNA repair inhibitors allowed for increased sensitivity in the detection of DNA breaks and gap intermediates during DNA repair. *The presence of breaks within the DM mutation, and expansion regardless of genomic DNA replication, suggest the involvement of DNA repair in the unstable properties of this mutation.*

There is a substantial variation in the repair rates in mammalian cells in the different stages in the cell cycle (Kaufmann and Wilson, 1990). Moreover, the efficiency of the

different DNA repair inhibitors varies substantially and is affected by the composition of the tissue culture media. For this reason the cells in the experiments that involved the detection of repair intermediates were submitted to a wide gradient of different repair inhibitors.

Gel electrophoresis analysis, together with enzymatic analysis, allowed partial differentiation between nicks and gaps (SSB) on one of the two strands, and DSBs. The results suggest that a high number of single-stranded nicks and gaps are present within the mutation although one cannot rule out the existence of DSBs. Those could be due to artefactual breakage of single-stranded DNA during extraction or handling, the presence of closely located nicks or gaps in complementary strands, and/or structure-based lesions induced *in situ.*

The causal relationship between the presence of breaks and the amplification of the DM mutation was confirmed. The presence of breaks is especially pronounced in cells in which the mutation is undergoing a more active amplification (data not shown).

A feature common to the trinucleotide repeat sequences causing human disease which have been described to date is that a high degree of instability is attained when the number of repeat units is sufficiently high. This instability is dependent on the length of the repeat sequence array (Chapter III). Interruptions of the perfect repeat, such as base substitutions, have been seen to render some of these repeat sequences more stable

(Eichler et al., 1994; Chung et al., 1993). *In vitro* studies show that CTG repeat arrays and their complementary CAG arrays form alternative DNA structures, such as hairpin-like structures (Gacy et al., 1995; Fry et al., 1994; Mitas et al., 1995; Pearson and Sinden, 1996; Mariappan et al., 1996). Peculiar to the DM mutation sequence are its DNA-histone interactions (Wang et al., 1994b; Wang and Griffith, 1995; Godde and Wolffe, 1996). Chromatin alterations have also been found in the vicinity of the DM mutation *in vivo* (Otten and Tapscott, 1995). These data are suggestive of structural transitions based on both the composition and length of the sequence.

In addition to these structural effects, the stability of this mutation must also be influenced by chromosomal location. The dynamics of the DM and other trinucleotide repeat mutations are usually quite different in different tissues (Chapter V), different diseases (even when sharing the same repeat sequence, Table 1-1), and especially in transgenic mouse models (Bingham et al., 1995; Burright et al., 1995).

Interestingly, in adult DM patients, the length of the CTG repeat array is significantly higher in skeletal muscle (primary target of DM pathology and highest expression of the DMK gene) than in leukocytes (Anvret et al., 1993; Thornton et al., 1994a; Zatz et al., 1995). This is consistent with the hypothesis that *DNA repair* is involved in the expansion of the DM mutation. DNA repair, at least nucleotide excision repair, is more rapid at or near transcriptionally active genes (Bohr et al., 1985; Hanawalt, 1989; Bohr, 1991). Long arrays of CTG repeats could become sequence-dependent pause sites for

RNA polymerase II and mistakenly trigger a repair reaction cascade. The frequency of repair would depend, in these conditions, on the transcriptional activity of the specific site.

Slipped-strand mispairing has been proposed as a mechanism for the evolution of simple repeat sequences and also as a cause of frameshift mutations in such sequences (Levinson and Gutman, 1987; Murphy et al., 1989). Instability of the DM mutation seems to occur (at least partially) by a sequence-length directed mechanism. There is no indication that any of the genes involved in microsatellite stability/instability are differentially affected in DM patients. Moreover, the normal size triplet repeats in DM and fragile X are stable in human mutator mismatch repair cell lines, which show general instability of microsatellites (Kramer et al., 1996).

Slippage by itself, as a cause of trinucleotide repeat expansion, cannot explain multiple findings, at least not in DM. For example, it cannot explain (a) the high instability acquired beyond a minimal threshold of repeating units, (b) the high tendency of the mutation to expand, rather than contract, (c) the variable outcome on mutation transmission according to the gender of the transmitting parent, (d) the apparent cell-type and not necessarily mitotic activity linked dynamics of the mutation, (e) the different behaviour of the mutation when in a different genomic context (transgenic models), and (f) the fact that there is clear instability whether or not there is cell-cycle-linked DNA replication. In contrast, DNA repair could explain all of these results.

Recognition and processing of DNA damage

There is a close relationship between the enzymes involved in DNA replication and DNA repair. The proteins involved have been seen to act in both processes, and damage recognition has been seen to be mediated through pauses in DNA and RNA polymerization. The pauses occur at sites perceived as lesions. DNA damage can be defined, in a cellular context, as any unusual alteration in the structure of this molecule. The cell has evolved pathways designed to deal with such altered structures by repairing them. Some of the replication and repair enzymes, eg, polymerases α, β, δ, and ε, have direct counterparts in yeast, whereas enzymes such as poly(ADP-ribose) polymerase, p53, and DNA-dependent protein kinase, which are elevated in response to DNA strand breaks, have not been found in the same organism. It appears, thus, that the cellular responses to DNA strand breaks differ considerably between higher and lower eukaryotes.

Excision repair is a cellular pathway involved in the processing of DNA damage. There are three main forms of DNA-excision-repair pathways, with little overlap between them: (a) nucleotide-excision repair, (b) mismatch repair, and (c) base-excision repair (Lindahl, 1995). Base-excision repair is a relatively well understood repair pathway.

Base-excision repair in mammalian cells involves recognition of damage at a specific base by a specific glycosylase which modifies the damaged base. The resulting information-less base site is then recognized by an endonuclease that catalyses the

hydrolysis of the base-sugar bond. The excised residue(s) result in repair patches of variable size. DNA polymerase β fills-in the gap (Dianov et al., 1992; Singhal et al., 1995). The damaged base(s) is displaced and its deoxyribose-phosphate residue is excised. The last step involves ligation (probably through ligase III/XRCC1), which restores the initial, undamaged DNA structure. The ligases involved seem to be distributed in a tissue specific fashion.

Signals at DNA strand breaks

Long branched polymers of poly (ADP-ribose) are synthesized very rapidly in response to DNA strand interruptions resulting from ionizing radiation or alkylation. The function of these polymers in response to DNA damage is unknown. It has been proposed that they may serve to reorganize the chromatin structure at the lesion site (Panzeter et al., 1993) in order to suppress homologous recombination in tandem repeat sequences which would result in genomic instability. PARP, which is associated with the nuclear matrix, is the enzyme responsible for this synthesis and has been seen to undergo extensive automodification following DNA damage.

Several other nuclear enzymes are known to compete for strand breaks in DNA. These include the large DNA-dependent protein kinase that is defective in *scid* cells (Blunt et al., 1995), and other protein kinases. The binding and activation of such enzymes at those breaks apparently leads to a stress response with the generation of signals that account for the accumulation of the p53 protein (Lu and Lane, 1993). The short half-life

of p53 depends on its degradation through a ubiquitin-dependent pathway. It has been proposed (Lindahl, 1995) that the accumulation and stabilization of p53 should be dependent on the interference of these DNA-break-response proteins with ubiquitin-conjugating enzymes. Supporting this idea is the observation that *S cerevisiae* mutants which are hypersensitive to DNA damage are also defective in ubiquitin conjugation (Jentsch et al., 1987).

A model to explain the instability of the DM mutation

DNA breaks are intermediates of DNA repair processes. The detection of breaks within the DM mutation, mutation expansion despite lack of chromosomal DNA replication, and other findings, make the hypothesis that DNA repair is the basis for the expansion of this mutation a compelling one. Although slippage could well be involved in this instability a simple model of slippage cannot explain multiple observations.

I propose a model of "macro-structure repair" to explain the dynamics of the DM mutation. In this model the DM mutation becomes a repair "hotspot" by virtue of its size, sequence composition, and location in a transcriptionally active region of the genome. The structure adopted by the long CTG (and complementary CAG) repeat arrays in the genomic region where the mutation is located becomes the primary signal for repair. This process could perhaps be induced directly by the recognition of the chromatin structure adopted by these long repetitive arrays by structure-specific enzymes (Harrington and Lieber, 1994) or indirectly by polymerase pauses. The

sequence itself could, by nature of its composition, present a challenge to the cellular repair mechanism. Amplification would then occur through active repair attempts and reiterative generation of repair patches at endonucleolytic nicks situated at the base of complementary strands in possible loops or hairpins. Sequence-induced error-prone repair DNA synthesis may in turn, perhaps due to slippage, increase the length of the repeat sequence array and in this way trigger the enzymatic machinery involved into a positive feedback mode. Slightly increased mutation size in cells that undergo active DNA replication, compared to quiescent cells (Figure 6-2), suggests a role for slippage in the described repair DNA synthesis.

It is tempting to imagine a common basis for the mechanism(s) by which instability of the different trinucleotide repeat sequences described in Table 1-1 occurs. Nevertheless, if as seen in some trinucleotide repeat disease mutations, the intergenerational dynamics depend on the genotype of the normal chromosome (Igarashi et al., 1996), a common mechanism would have to allow for this possibility.

The gene conversion seen in the DM and fragile X regions (van den Ouweland et al., 1994) may be due to the presence of DSBs, since there appears to be a correlation between the two processes (Orr-Weaver et al., 1988; Weng et al., 1996). It may be significant that there seems to be a hotspot for deletions in the CGG repeat region of FMR1 in fragile X patients (de Graaff et al., 1995). If randomly induced DSB occur in cis close to and/or within the mutation (due perhaps to torsional stress in this

chromosomal region) CTG repeat amplification could instead be the result of a break-healing process through breakage-fusion-bridge cycles.

Although little is known about them, post-replicative mechanisms are likely to operate in the elimination and/or repair of large DNA secondary structures (Leach, 1994). Pre- or post-replicative repair is likely to act as means of reducing the risk of introducing errors during the replication of damaged DNA (Kaufmann, 1989). Recognition of unusual large structures adopted by specific DNA sequences, may be more likely than, for example, single-base damage recognition and therefore not directly dependent on replication fork progression pauses or transcription.

Instead of being recognized through direct features of the DNA structure adopted by the long arrays of CTG repeats, the DM mutation may be recognized through the proteins that bind to this sequence, and this way become a substrate for nuclear enzymes involved in DNA metabolism. The DM mutation DNA sequence has been seen to bind to ssDNA binding proteins (Yano-Yanagisawa et al., 1995) and DNA mismatch repair enzymes (C.E. Pearson -Dep. of Biochem. & Biophys., Texas A & M University-personal communication). These proteins are also part of the DNA repair machinery.

Studies on the behaviour of the DM mutation out of its natural structural and environmental context, e.g., in *E coli* plasmids (Kang et al., 1996), may be of limited value for explaining its mechanism of expansion. The available data show instability

to be strongly modulated by the genetic and cellular environment although, because of the sequence effects on the local DNA topology, chromatin structure and stochastic components may also have major roles. Future research should focus on the study of the proteins that recognize, bind, and participate in the events leading to the unusual properties of this highly unstable sequence of DNA.

CONCLUSIONS

The epidemiological and molecular studies presented in this thesis provide evidence of the pivotal role of the size of the DM mutation in its own intergenerational dynamics. Moreover, the relative intergenerational stability of the DM protomutation suggests a higher frequency of mutant DM alleles in the general population than previously estimated, and provides a means of explaining the maintenance of DM in the population.

The intergenerational dynamics of this mutation are linked to both the number of CTG repeats and the gender of the transmitting parent. In general, there is a certain degree of predictability of the outcome of transmission of a mutant allele that depends on a combination of mutation size and whether it is a maternal or a paternal transmission.

CDM is a clinical form of the disease that is strongly influenced by the gender of the transmitting parent as well as by the number of CTG repeats in the mutant allele in both the mother and the fetus. A very high number of repeats in the offspring is neither a necessary nor a sufficient condition for CDM to occur. CDM also depends on a high number of repeats in the mutant maternal allele. The additive effect of both suggests that the distinctive phenotypic features of CDM are only indirectly derived from the presence of an expanded CTG repeat sequence in the cells.

The study of the size of the mutant allele(s) in sperm cells in individuals with the disease revealed comparatively greater expansions in this tissue than in peripheral blood leukocytes. The size of the mutant alleles in the offspring of these individuals revealed that pre-zygotic events are important factors in the intergenerational dynamics of the DM CTG repeat.

In general, the tissues from DM patients showed increased somatic mosaicism and increased mutation size and heterogeneity as a function of time. These increases were also evident in patients' cells maintained in tissue culture.

In addition, the CTG repeat expansion was found to take place in the absence of cell-cycle-coupled DNA replication. Separate studies revealed that recombination between homologous chromosomes and effects of the normal allele on the mutant allele could be discarded as major sources of instability. This, together with the results of studies of somatic and intergenerational instability of the mutation, left *DNA repair* as a prime candidate to be explored. The presence of repair intermediates exclusive to this mutation, together with findings concerning the mutation's structure and dynamics, suggest that DNA repair plays a major role in causing the high instability of this sequence. The observation of a slightly higher expansion in actively cycling cells, compared to mitotically inactive cells, may also suggest a role for slippage within the repair processes of the DM mutation.

REFERENCES

Abbruzzese, C., Krahe, R., Liguori, M., Tessarolo, D., Siciliano, M.J., Ashizawa, T. and Giacanelli, M. (1996). Myotonic dystrophy phenotype without expansion of (CTG)n repeat - an entity distinct from proximal myotonic dystrophy. *J of Neurol* 243, 715-721

Abeliovich, D., Lerer, I., Pashut-Lavon, E., Shmueli, E., Rass-Rothschild, A. and Frydman, M. (1993). Negative expansion of the myotonic dystrophy unstable sequence. *Am J Hum Genet* 52, 1175-1181

Anvret, M., Åhlberg, G., Grandell, U., Hedberg, B., Johnson, K. and Edström, L. (1993). Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy. *Hum Molec Genet* 2, 1397-1400

Armour, J.A.L., Patel, I., Thein, S.L., Fey, M.F., and Jeffreys, A.J. (1989). Analysis of somatic mutations at human minisatellite loci in tumours and cell lines. *Genomics* 4, 328-334

Ashizawa, T., Anvret, M., Baignet, M., **Barceló, J.M.**, Brunner, H., Cobo, A.M., Dallapicola, B., Fenwick, R.G., Grandell, U., Harley, H., Junien, C., Koch, M.C., Korneluk, R.G., Lavedan, C., Miki, T., Mulley, J.C., López de Munain, A., Novelli, G., Roses, A.D., Seltzer, W.K., Shaw, D.J., Smeets, H., Sutherland, G.R., Yamagata, H. and Harper, P.S. (1994). Characteristics of intergenerational contractions of the CTG repeat in myotonic dystrophy. *Am J Hum Genet.* 54, 414-423

Aslanidis, C., Jansen, G., Amemiya, C., Shutler, G., Mahadevan, M., Tsilfidis, C., Chen, C., Alleman, J., Wormskamp, N.G.M., Vooijs, M., Buxton, J., Johnson, K., Smeets, H.J.M., Lennon, G.G., Carrano, A.V., Korneluk, R.G., Wieringa, B. and de Jong, P. (1992). Cloning of the essential myotonic dystrophy region and mapping of the putative defect. *Nature* 355, 548-551

Barceló, J.M., Mahadevan, M.S, Tsilfidis, C. MacKenzie, A.E. and Korneluk, R.G. (1993). Intergenerational stability of the myotonic dystrophy protomutation. *Hum Molec Genet* 2, 705-709

Barceló, J.M., Pluscauskas, M., MacKenzie, A.E., Tsilfidis, C., Narang, M. and Korneluk, R.G. (1994). Additive influence of maternal and offspring DM-Kinase gene CTG repeat lengths in the genesis of congenital myotonic dystrophy. *Am J Hum Genet* 54, 1124-1125

Barceló, J.M. and Korneluk, R.G. (1997). Cell-cycle-linked DNA replication independent mechanism of expansion of the DM mutation: assessment of DNA repair processes. Submitted manuscript.

Biancalana, V., Serville, F., Pommier, J., Julien, J., Hanauer, A., and Mandel, J.L. (1992). Moderate instability of the trinucleotide repeat in spino bulbar muscular atrophy. *Hum Molec Genet* 1, 255-258

Bingham, P.M., Scott, M.O., Wang, S., McPhaul, M.J., Wilson, E.M., Garbern, J.Y., Merry, D.E. and Fischbeck, K.H. (1995). Stability of an expanded trinucleotide repeat in the androgen receptor gene in transgenic mice. *Nat Genet* 9, 191-196

Birnboim, H.C. (1992). Extraction of high molecular weight RNA and DNA from cultured mammalian cells. *Methods Enzymol* 216, 154-160

Blunt, T., Finnie, N.J., Taccioli, G.E., Smith, G.C.M., Demengeot, J., Gottlieb, T.M., Mizuta, R., Varghese, A.J., Alt, F.W., Jeggo, P.A. and Jackson, S.P. (1995). Defective DNA-dependent protein kinase activity is linked to V(D)J recombination and DNA repair defects associated with the murine *scid* mutation. *Cell* 80, 1-20

Bohr, V.A., Smith, C.A., Okumoto, D.S. and Hanawalt, P.C. (1985). DNA repair in an active gene: removal of pyrimidine dimers from the DHFR gene of CHO cells is much more efficient than in the genome overall. *Cell* 40, 359-369

Bohr, V. (1991). Gene specific DNA repair. *Carcinogenesis* 12, 1983-1992

Boucher, C.A, King, S.K., Carey, N., Krahe, R., Winchester, L., Rahman, S., Creavin, T, Meghji, P., Bailey, M.E., Chartier, F.L., Brown, S.D., Siciliano, and M.J., Johnson, K.J. (1995). A novel homeodomain-encoding gene is associated with a large CpG island interrupted by the myotonic dystrophy unstable (CTG)n repeat. *Hum Molec Genet* 4, 1919-1925

Bronner, C.E., Baker, S.M., Morrison, P.T., Warren, G., Smith, L.G., Lescoe, M.K., Kane, M., Earabino, C., Lipford, J., Lindblom, A., Tannergard, P., Bollag, R.J., Godwin, A.R., Ward, D.C., Nordenskjold, M., Fishel, R., Kolodner, R. and Liskay, M. (1994). Mutation in the DNA mismatch repair gene homologue *hMLH1* is associated with hereditary non-polyposis colon cancer. *Nature* 368, 258-261

Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Abutarani, H., Hunter, K., Stanton, V.P., Thirion, J.-P., Hudson, T., Sohn, R., Zemelman, B., Snell, R.G., Rundle, S.A., Crow, S., Davies, J., Shelbourne, P., Buxton, J., Jones, C., Juvonen, V., Johnson, K., Harper, P.S., Shaw, D.J. and Housman, D.E. (1992). 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* 68, 799-808

Brown, W.T., Houck, G.E. Jr, Jeziorowska, A., Levinson, F.N., Ding, X., Dobkin, C., Zong, N., Henderson, J., Brooks, S.S. and Jenkins, E.C. (1993). Rapid fragile X carrier screening and prenatal diagnosis by a non-radioactive PCR test. *JAMA* 270, 1569-1575

Brunner, H.G., Jansen, G., Nillesen, W., Nelen, M.R., de Die, C.E.M., Howeler, C.J., van Oost, B.A., Wieringa, B., Ropers, H-H. and Smeets, H.J.M. (1993a). Brief report: Reverse mutation in myotonic dystrophy. *The New England J of Med* 238, 476-480

Brunner, H.G., Bruggenwirth, H.T., Nillesen, W., Jansen, G., Hamel, B.C., Hoppe, R.L., de Die, C.E., Howeler, C.J., van Oost, A. and Wieringa, B. (1993b). Influence of the transmitting parent as well as of parental allele size on the CTG expansion in myotonic dystrophy (DM). *Am J Hum Genet* 53, 1016-1023

Burright, E.N., Clark, H.B., Servadio, A., Matilla, T., Feddersen, R.M., Yunis, W.S., Duvick, L.A., Zoghbi, H.Y. and Orr, H.T. (1995). SCA1 transgenic mice: a model for neurodegeneration caused by an expanded CAG trinucleotide repeat. *Cell* 82, 937-948

Buxton, J., Shelbourne, P., Davies, J., Jones, C., Van Tongeren, T., Aslanidis, C., de Jong, P., Jansen, G., Anvret, M., Riley, B., Williamson, R. and Johnson, K. (1992). Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* 355, 547-548

Campuzano, V., Montermini, L., Moltò, M.D., Pianese, L., Cossée, M., Cavalcanti, F., Monros, E., Rodius, F., Duclos, F., Monticelli, A., Zara, F., Canizares, J., Koutnikova, H., Bidichandani, S.I., Gellera, C., Brice, A., Trouillas, P., De Michele, G., Filla, A., De Frutos, R., Palau, F., Patel, P.I., Di Donato, S., Mandel, J.-L., Cocozza, S., Koening, M., and Pandolfo, M. (1996). Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* 271, 1423-1427

Carango, P., Noble, J.E., Marks, H.G., and Funanage, V.L. (1993). Absence of myotonic dystrophy protein kinase (DMPK) mRNA as a result of a triplet repeat expansion in myotonic dystrophy. *Genomics* 18, 340-348

Caskey, C.T., Pizzuti, A., Fu, Y-H, Fenwick, R.G., and Nelson, D.L. (1992). Triple repeat mutations in human disease. *Science* 256, 784-788

Cheng, S., **Barceló, J.M.**, and Korneluk, R.G. (1996). Characterization of large CTG repeat expansions in myotonic dystrophy alleles using PCR. *Hum Mutat* 7, 304-310.

Chung, M.-Y., Ranumm, L.P.W., Duvick, L.A., Servadio, A., Zoghbi, H.Y., and Orr, H.T. (1993). Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. *Nat Genet* 5, 254-258

David, G., Giunti, P., Abbas, N., Coullin, P., Stevanin, G., Horta, W., Gemmill, R., Weissenbach, J., Wood, N., Cunha, S., Drabkin, H., Harding, A.E., Agid, Y., and Brice, A. (1996). The gene for autosomal dominant cerebellar ataxia Type II is located in a 5-cM Region in 3p12-13: genetic and physical mapping of the SCA7 locus. *Am J Hum Genet* 59, 1328-1336

Davies, J., Yamagata, H., Shelbourne, P., Buxton, J., Ogihara, T., Nokelainen, P., Nakagawa, M., Williamson, R., Johnson, K., and Miki, T. (1992). Comparison of the myotonic dystrophy associated CTG repeat in European and Japanese populations. *J Med Genet* 29, 766-769

de Graaff, E., Rouillard, P., Williems, P.J., Smits, A.P.T., Rousseau, F., and Oostra, B.A. (1995). Hotspot for deletions in the CGG repeat region of *FMR1* in fragile X patients. *Hum Molec Genet* 4, 45-49

Devys, D., Biancalana, V., Rousseau, F., Boue, J., Mandel, J.L., and Oberlé, I. (1992). Analysis of full fragile X mutations in fetal tissues and monozygotic twins indicate that abnormal methylation and somatic heterogeneity are established in early development. *Am J Med Genet* 43, 208-216

Dianov, G., Price, A., and Lindahl, T. (1992). Generation of single-nucleotide repair patches following excision of uracil residues from DNA. *Mol Cell Biol* 12, 1605-1612

Dib, C., Fauré, S., Fizames, C., Samson, D., Drouot, N., Vigal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J., and Weissenbach, J. (1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380, 152-154

Drake, J.W., Glickman, B.W., and Ripley, L.S. (1983). Updating the theory of mutation. *Am Scientist* 71, 621-630

Dunne, P.W., Ma, L., Caskey, D.L., Harati, Y., and Epstein, H.F. (1996). Localization of myotonic dystrophy protein kinase in skeletal muscle and its alteration with disease. *Cell Motil and the Cytoskel* 33, 52-63

Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M., O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G., Weissman, S.W., Slightom, Blechl, A.E., Smithies, O., Baralle, F.E., Shoulders, C.C., and Proudfoot, N.J. (1980). The structure and evolution of the human beta-globin gene family. *Cell* 21, 653-668

Eichler, E.E., Holden, J.J., Popovich, B.W., Reiss, A.L., Snow, K., Thibodeau, S.N., Richards, C.S., Ward, P.A., and Nelson, D.L. (1994). Length of uninterrupted CGG repeats determines instability of the FMR1 gene. *Nat Genet* 8, 88-94

Feinberg, A.P. and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* 132, 6-13

Fisch, G.S., Snow, K., Thibodeau, S.N., Chalifaux, M., Holden, J.J., Nelson, D.L., Howard-Peebles, P.N., and Maddalena, A. (1995). The fragile X premutation in carriers and its effect on mutation size in offspring. *Am J Hum Genet* 56, 1147-1155

Fishel, R., Lescoe, M.K., Rao, M.R., Copeland, N.G., Jenkins, N.A., Garber, J., Kane, M., and Kolodner. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* 75, 1027-1038

Fry, M. and Loeb, L.A. (1994). The fragile X syndrome d(CGG)n nucleotide repeats form a stable tetrahelical structure *Proc Natl Acad Sci USA* 91, 4950-4954

Fu Y.-H., Kuhl, D.P.A., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkerk, A.J.M.H., Holden, J.J.A., Fenwick Jr, R.G., Warren, S.T., Oostra, B.A., Nelson, D.L. and Caskey, C.T. (1991). Variation of the CGG repeat at the Fragile-X site results in genetic instability: resolution of the Sherman paradox. *Cell* 67, 1047-1058

Fu, Y.-H., Pizzuti, A., Fenwick Jr, R.G., King, J., Rajnarayan, S., Dunne, P.W., Dubel, J., Nasser, G.A., Ashizawa, T., de Jong, P., Wieringa, B., Korneluk, R., Perryman, M.B., Epstein, H.F., and Caskey, C.T. (1992). An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* 255, 1256-1258

Fu, Y.-H., Friedman, D.L., Richards, S., Pearlman, J.A., Gibbs, J.A., Pizzuti, A., Ashizawa, A., Perryman, M.B., Scarlato, G., Fenwick, R.G.Jr., and Caskey, C.T. (1993). Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* 260, 235-238

Gacy, A.M., Goellner, G., Juranic, N., Macura, S., and McMurray, C.T. (1995). Trinucleotide repeats that expand in human disease form hairpin structures in vitro. *Cell* 18, 533-540

Glover, T.W., Berger, C., Coyle J., and Echo, B. (1984). DNA polymerase α inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. *Hum Genet* 67, 136-142

Godde, J.S. and Wolffe, A.P. (1996). Nucleosome assembly on CTG triplet repeats. *J Biol Chem* 271, 15222-15229

Goulian, M., Richards, S.H., Heard, C.J. and Bigsby, B.M. (1990). Discontinuous DNA synthesis by purified mammalian proteins. *J Biol Chem* 265, 18461-18471

Griggs, R.C. and Wood D.S. (1989). Criteria for establishing the validity of genetic recombination in myotonic dystrophy. *Neurology* 39, 420-421

Grimaldi, K.A., Bingham, J.P., Souhami, R.L. and Hartley, J.A. (1994). DNA damage by anticancer agents and its repair: mapping in cells at the subgene level with quantitative polymerase chain reaction. *Anal Biochem* 222, 236-242

Hamada, H., Petrino, M., and Kakunaga T. (1982). A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomics. *Proc Natl Acad Sci USA* 79, 6465-6469

Hamshere M.G. and Brook J.D. (1996). Myotonic dystrophy, knockouts, warts and all. *TIG* 12, 332-334

Hanawalt, P.C. (1989). Preferential repair of damage in actively transcribed DNA sequences in vivo. *Genome* 31, 605-611

Harley, H.G., Runde, S.A., Reardon, W., Myring, J., Crow, S., Brook, J.D., Harper, P.S., and Shaw, D.J. (1992a). Unstable DNA sequence in myotonic dystrophy. *The Lancet* 339, 1125-1128

Harley, H.G., Brook, J.D., Rundle, S.A., Crow, S., Reardon, W., Buckler, A.J., Harper, P.S., Housman, D.E., and Shaw, D.J. (1992b). Expansion of an unstable DNA region and phenotypic variation in myotonic dystrophy. *Nature* 355, 545-546

Harley, H.G., Rundle, S.A., MacMillan, J.C., Myring, J., Brook, J.D., Crow, S., Reardon W., Fenton, I., Shaw, D.J., and Harper, P.S. (1993). Size of the unstable CTG repeat sequence in relation to phenotype and parental transmission of myotonic dystrophy. *Am J Hum Genet* 52, 1164-1174

Harper, P.S. (1975). Congenital myotonic dystrophy in Britain. *Arch Dis Child* 50, 505-513

Harper, P.S. (1989). *Myotonic Dystrophy*, 2nd edn., London and Philadelphia, Saunders.

Harrington, J.J. and Lieber, M.R. (1994). The characterization of a mammalian DNA structure-specific endonuclease. *EMBO J* 13, 1235-1246

Harris, S., Moncrieff, C., and Johnson, K. (1996). Myotonic dystrophy: will the real gene please step forward! *Hum Molec Genet* 5, 1417-1423

Hendrich, B.D. and Willard, H.F. (1995). Epigenetic regulation of gene expression: the effect of altered chromatin structure from yeast to mammals. *Hum Molec Genet* 4, 1765-1777

Hecht, B.K., Donnelly, A., Gedeon, A.K., Byard, R.W., Hann, E.A., and Mulley, J.C. (1993). Direct molecular diagnosis of myotonic dystrophy. *Clin Genet* 43, 276-285

Hoffman-Radvanyi, H., Lavedan, C., Rabes, J.P., Savoy, D., Duros, C., Johnson, K., and Junien, C. (1993). Myotonic dystrophy: absence of CTG enlarged transcript in congenital forms, and low expression of the normal allele. *Hum Molec Genet* 2, 1263-1266

Howeler, C.J., Busch, H.F.M., Geraedts, J.P.M., Niermeijer, M.F., and Staal, A. (1989). Anticipation in myotonic dystrophy: fact or fiction? *Brain* 112, 779-797

Hunter, A., Tsilfidis, C., Mettler, G., Jacob, P., Mahadevan, M., Surh, L. and Korneluk, R.G. (1992). The correlation of age of onset with CTG trinucleotide repeat amplification in myotonic dystrophy. *J of Med Genet* 29, 774-779

Huselbos, T., Brunner, H., Wieringa, B., Friedrich, U., Smeets, D., Oei, T., Hustinx, T., Scheres, J., Wienker, T., Humphries, S., Myklebost, O., Junien, C., Haarter, B., and Ropers, H.H. (1985). Regional assignment of C3, GP1, ApoC2 and β-HCG and their linkage relationships with DM and 19 cen. *Cytogenet Cell Genet* 40, 658

Igarashi, S., Takiyama, Y., Cancel, G., Rogaeva, E.A., Sasaki, H., Wakisaka, A., Zhou, Y.X., Takano, H., Endo, K., Sanpei, K., Oyake, M., Tanaka, H., Stevanin, G., Abbas, N., Durr, A., Rogaev, E.I., Sherrington, R., Tsuda, T., Ikeda, M., Cassa, E., Nishizawa, M., Benomar, A., Julien, J., Weissenbach, J., Wang, G.X., Agid, Y., St George-Hyslop, P.H., Brice, A., and Tsuji, S. (1996). Intergenerational instability of the CAG repeat of the gene for Machado-Joseph disease (*MJD1*) is affected by the genotype of the normal chromosome: implications for the molecular mechanism of the instability of the CAG repeat. *Hum Mol Genet* 5, 923-932

Imbert, G., Saudou, F. Yvert, G., Devys, D., Trotier, Y., Garnier, J.M., Weber, C., Mandel, J.L., Cancel, G., Abbas, N., Durr, A., Didierjean, O., Stevanin, G., Agid, Y., and Brice, A. (1996). Cloning the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nat Genet* 14, 285-292

Ishimi, Y., Claude, A., Bullock, P., and Hurwith, J. (1988). Complete enzymatic synthesis of DNA containing the SV40 origin of replication. *J Biol Chem* 263, 19723-19733

Jansen, G., de Jong, P.J., Anemiya, C., Aslanidis, C., Shaw, D.J., Harley, H.G., Brook, J.D., Fenwick, R., Korneluk, R.G., Tsilfidis, C., Shutler, G., Hermens, R., Wormskamp, N.G.M., Smeets, H.J., and Wieringa, B. (1992a). Physical and genetic characterization of the distal segment of the myotonic dystrophy area on 19q. *Genomics* 13, 509-517

Jansen, G., Mahadevan, M., Amemiya, C., Wormskamp, N., Segers, B., Hendricks, W., O'Hoy, K., Baird, S., Sabourin, L., Lennon, G., Jap, P.L., Iles, D., Coerwinkel, M., Hofker, M., Carrano, A.V., de Jong, P.J., Korneluk, R.G., and Wieringa, B. (1992b). Characterization of the myotonic dystrophy region predicts multiple protein isoform-encoding mRNAs. *Nat Genet* 1, 261-266

Jansen, G., Bartolomei, M., Kalscheuer, V., Merkx, G., Wormskamp, N., Mariman, D., Smeets, D., Ropers, H.H., and Wieringa, B. (1993). No imprinting involved in the expression of DM-kinase mRNAs in mouse and human tissues. *Hum Molec Genet* 2, 1221-1227

Jansen, G., Bachner, D., Coerwinkel, M., Wormskamp, N., Hameister, H., and Wieringa, B. (1995). Structural organization and developmental expression pattern of the mouse WD-repeat gene DMR-N9 immediately upstream of the myotonic dystrophy locus. *Hum Mol Genet* 4, 843-852

Jansen, G., Groenen P.J., Bachner, D., Jap, P.H., Coerwinkel, M., Oerlemans, F., van den Broel W., Gohlsh, B., Pette, B., Plomp, J.J., Molenaar, P.C., Nederhoff, M.G., van Echteld, C.J., Dekker, M., Berns, A., Hameister, H., and Wieringa, B. (1996). *Nat Genet* 13, 316-324

Jeffreys, A.J., Wilson, V., and Thein, S.L. (1985). Hypervariable 'minisatellite' regions in human DNA. *Nature* 314, 67-73

Jeffreys, A.J., Royle, N.J., Wilson, V., and Wong, Z. (1988). Spontaneous mutation rates to new length alleles at tandem-repetitive hypervariable loci in human DNA. *Nature* 332, 278-281

Jentsch, S., McGrath, J.P., and Varshavsky, A. (1987). The yeast DNA repair gene RAD6 encodes a ubiquitin-conjugated enzyme. *Nature* 329, 131-134

Jodice, C., Malaspina, P., Persichetti, F., Novelletto, A., Spandaro, M., Giunti, P., Morocutti, C., Terrenato, L., Harding, A.E., and Frontali, M. (1994). Effect of trinucleotide repeat length and parental sex on phenotypic variation in spinocerebellar ataxia I. *Am J Hum Genet* 54, 959-965

Kang, S., Ohshima, K., Jaworski, A., and Wells, R.D. (1996). CTG triplet repeats from the myotonic dystrophy gene are expanded in Escherichia coli distal to the replication origin as a single large event. *J Mol Biol* 258, 543-547

Kaufmann, W.K. (1989). Pathways of human cell post-replication repair. *Carcinogenesis* 10, 1-11

Kaufmann, W.K. and Wilson, S.J. (1990). DNA repair endonuclease activity during synchronous growth of diploid human fibroblasts. *Mutat Res* 236, 107-117

Kelly, R., Bulfield, G., Collick, A., Gibbs, M., and Jeffryes, A.J. (1989). Characterization of a highly unstable mouse minisatellite locus: evidence for somatic mutation during early development. *Genomics* 5, 844-856

Koch, M.C., Grimm, T., Harley, H.G. and Harper, P.S. (1991). Genetic risks for children of women with myotonic dystrophy. *Am J Hum Genet* 48, 1084-1091

Koga, R., Nakao, Y., Kurano, Y., Tsukahara, T., Nakamura, A., Ishiura, S., Nonaka, I., and Arahata, K. (1994). Decreased myotonin-protein kinase in the skeletal and cardiac muscles in myotonic dystrophy. *Biochem Biophys Res Commun* 202, 577-585

Krahe, R., Ashizawa, T., Abbruzzese, C., Roeder, E., Carango, P., Giacanelli, M., Funanage, V.L., and Siciliano, M.J. (1995a). Effect of myotonic dystrophy trinucleotide repeat expansion on DMPK transcription and processing. *Genomics* 28, 1-14

Krahe, R., Eckhart, M., Ogunniyi, A.O., Osuntokun, B.O., Siciliano, M.J., and Ashizawa, T. (1995b). De novo myotonic dystrophy mutation in a Nigerian kindred. *Am J Hum Genet* 56, 1067-1074

Kramer, P.R., Pearson, C.E., and Sinden, R.R. (1996). Stability of triplet repeats of myotonic dystrophy and fragile X loci in human mutator mismatch repair cell lines. *Hum Genet* 98, 151-157

Krawczak, M. and Cooper, D.N. (1991). Gene deletions causing human genetic disease: mechanisms of mutagenesis and the role of the local DNA sequence environment. *Hum Genet* 86, 425-441

Kremer, E.J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S.T., Schlessinger, D., Sutherland, G.R., and Richards, R.I. (1991). Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science* 252, 1711-1714

Kremer, B., Almqvist, E., Theilmann, J., Spence, N., Telenius, H., Goldberg, Y.P., and Hayden, M.R. (1995). Sex-dependent mechanisms for expansions and contractions of the CAG repeat on affected Huntington disease chromosomes. *Am J Hum Genet* 57, 343-350

Kunkel, T.A., Roberts, J.D., Thomas, D.C., and Nguyen, D.C. (1992). "The fine structure" of DNA replication fidelity. *Alfred Benzon Symposium* 35, 189-199

La Spada, A.R., Wilson, E.M., Lubahn, D.B., Harding, A.E., and Fischbeck, K.H. (1991). Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* 352, 77-79

Lavedan, C., Hofmann-Radvanyi, H., Shelbourne, P., Rabes, J.-P., Duros, C., Savoy, D., Dehaupas, I., Luce, S., Johnson, K., and Junien, C. (1993a). Myotonic dystrophy: Size- and sex-dependent dynamics of CTG meiotic instability, and somatic mosaicism. *Am J Hum Genet* 52, 875-883

Lavedan, C., Hofmann-Radvanyi, H., Rabes, J.P., Roume, J., and Junien, C. (1993b). Different sex dependent constraints in CTG length variation as explanation for congenital myotonic dystrophy. *The Lancet* 341, 237

Leach, D.R.F. (1994). Long DNA palindromes, cruciform structures, genetic instability and secondary structure repair. *Bioessays* 16, 893-900

Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L.A., and Nystrom-Lahti, M. (1993). Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell* 75, 1215-1225

Levinson, G. and Gutman, G.A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Molec Biol Evol* 4, 203-221

Lindahl, T. (1995). Recognition and processing of DNA damage. *J of Cell Science*, Supplement 19, 73-77

Litt, M. and Lutty, J.A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a direct dinucleotide repeat within the cardiac muscle actin gene. *Am J Hum Genet* 44, 397-401

Loesch, D.Z., Huggins, R., Petrovic, V., and Slater, H. (1995). Expansion of the CGG repeat in Fragile X in the FMR1 gene depends on the sex of the offspring. *Am J Hum Genet* 57, 1408-1413

Lu, X. and Lane, P.D. (1993). Differential induction of transcriptionally active p53 following UV or ionizing radiation: Defects in chromosome instability syndromes? *Cell* 75, 765-778

MacKenzie, A.E., MacLeod, H.L., Hunter, A.G.H., and Korneluk, R.G. (1989). Linkage analysis of the apolipoprotein C2 gene and myotonic dystrophy on human chromosome 19 reveals linkage disequilibrium in a French-Canadian population. *Am J Hum Genet* 44, 140-147

Maeda, M., Taft, C.S., Bush, E.W., Holder, E., Bailey, W.M., Neville, H., Perryman, M.B., and Bies, R.D. (1995). Identification, tissue-specific expression, and subcellular localization of the 80- and 71-kDa forms of myotonic dystrophy kinase protein. *J Biol Chem* 270, 20246-20249

Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Anemiya, C., Jansen, G., Neville, C., Narang, M., **Barceló, J.**, O'Hoy, K., Leblond, S., Earle-MacDonald, J., de Jong, P., Wieringa, B., and Korneluk, R.G. (1992). Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* 255, 1253-1255

Mahadevan, M.S., Foitzik, M.A., Surh, L.C., and Korneluk, R.G. (1993a). Characterization and polymerase chain reaction (PCR) of an Alu deletion polymorphism on total linkage disequilibrium with myotonic dystrophy. *Genomics* 15, 446-448

Mahadevan, M.S., Amemiya, C., Jansen, G., Sabourin, L., Baird, S., Neville, C.E., Wormskamp, N., Segers, B., Batzer, M., Lamerdin, J., de Jong, P., Wieringa, B., and Korneluk, R.G. (1993b). Structure and genomic sequence of the myotonic dystrophy (DM kinase) gene. *Hum Mol Genet* 2, 299-304

Mariappan, S.V., Catasti, P., Chen, X., Ratliff, R., Moyzis, R.K., Bradbury, E.M., and Gupta, G. (1996). Solution structures of the individual single strands of the fragile X DNA triplets (GCC)n.(GGC)n. *Nucleic Acids Res* 24, 784-792

Mathieu, J., De Braekeleer, M., and Provost, C. (1990). Genealogical reconstruction of myotonic dystrophy in the Saguenay-Lac-Saint-Jean area (Quebec, Canada). *Neurology* 40, 838-842

Miesfeld, R., Krystal, M., and Arnheim, N. (1981). A member of a new repeat sequence family which is conserved throughout eukaryotic evolution is found between the human γ and β globin genes. *Nucleic Acids Res* 9: 5931-5947

Milot, E., Fraser, P., and Grosveld, F. (1996). Position effect and genetic disease. *TIG* 12, 123-126

Mitas, M., Yu, A., Dill, J., Kamp, T.J., Chambers, E.J., and Haworth, I.S. (1995). Hairpin properties of single-stranded DNA containing a GC-rich triplet repeat: $(CTG)_{15}$. *Nucleic Acids Res* 23, 1050-1059

Monckton, D.G., Wong, L.J., Ashizawa, T., and Caskey, C.T. (1995). Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males: small pool PCR analyses. *Hum Mol Genet* 4, 1-8

Morgan, A.R., Coulter, M.B., Flintoff, W.F., and Paetkau, V.H. (1974). Enzymatic synthesis of deoxyribonucleic acids with repeating sequences. A new repeating trinucleotide deoxyribonucleic acid, d(T-C-C)$_n$.d(G-G-A)$_n$. *Biochemistry* 13, 1596-1603

Mulley, J.C, Staples, A., Donnelly, A., Gedeon, A.K, Hecht, B.K, Nicholson, G.A, Haan E.A., and Sutherland G.R. (1993). Explanation for exclusive maternal origin for congenital form of myotonic dystrophy. *The Lancet* 341, 236-237

Murphy, G.L., Connell, T.D., Barrit, D.S., Koomey, M., and Cannon, J.G. (1989). Phase variation of gonococcal protein II: regulation of gene expression by slipped-strand mispairing of a repetitive DNA sequence. *Cell* 56, 539-347

Myklebost, O., and Rogne, S. (1988). A physical map of the apolipoprotein gene cluster on human chromosome 19. *Hum Genet* 78, 244-247

Naegeli, H. (1994). Roadblocks and detours during DNA replication: mechanisms of mutagenesis in mammalian cells. *Bioessays* 16, 557-564

Nakagawa, M., Yamada, H., Higuchi, Y., Kaminishi, Y., Miki, T., Johnson, K., and Osame, M. (1994). A case of paternally inherited congenital myotonic dystrophy. *J Med Genet* 31, 397-400

Nakamura, Y., Lepppert, M., O'Connell, P., Wolfe, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., and Kumlin, E. (1987). Variable number of tandem repeats (VNTR) markers for human gene mapping. *Science* 235, 1616-1622

Nakamura, Y., Lathrop, M., O'Connell, P., Leppert, M., Barker, D., Wright, E., Skolnick, M., Kondoleon, S., Litt, M., and Lalouel, J.M. (1988). A mapped set of DNA markers for human chromosome 17. *Genomics* 2, 302-309

Neville, C.E., Mahadevan M.S., **Barceló, J.M.**, and Korneluk, R.G. (1994). High resolution genetic analysis suggests one ancestral predisposing haplotype for the origin of the myotonic dystrophy mutation. *Hum Molec Genet* 3, 45-51

Novelli, G., Gennarelli, M., Zelano, G., Pizzuti, A., Fattorini, C., Caskey, C.T., and Dallapiccola, B. (1993). Failure in detecting mRNA transcripts from the mutated allele in myotonic dystrophy muscle. *Biochem Mol Biol Int* 29, 291-297

Oberlé, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boué, J., Bertheas, M.F., and Mandel, J.L. (1991). Instability of a 550-base pair DNA segment and abnormal methylation in fragile-X syndrome. *Science* 252, 1097-1102

Ohya, K., Tachi, N., Chiba, S., Sato, T., Kon, S., Kikuchi, K., Imamura, S., Yamagata, H., and Miki, T. (1994). Congenital myotonic dystrophy transmitted from an asymptomatic father with a DM-specific gene. *Neurology* 44, 1958-1960

O'Hoy, K.L., Tsilfidis, C., Mahadevan, M.S., Neville, C.E., **Barceló, J.**, Hunter, A.G.W., and Korneluk, R.G. (1993). Reduction in size of the myotonic dystrophy trinucleotide repeat mutation during transmission. *Science* 259, 809-812

Orr-Weaver, T.L., Nicolas, A., and Szostak, J.W. (1988). Gene conversion adjacent to regions of double-strand break repair. *Mol Cell Biol* 8, 5292-5298

Otten A.D. and Tapscott, S.J. (1995). Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proc Natl Acad Sci USA* 92, 5465-5469

Panzeter, P.L., Zweifel, B., Malanga, M., Waser, S.H., Richard, M.-C. and Althaus, F.R. (1993). Targeting of histone tails by poly(ADP-ribose). *J Biol Chem* 268, 17662-17664

Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Hasaltine, W.A., Fleischmann, R.D., Fraser, C.M., and Adams, M.D. (1994). Mutation of a mutL homolog in hereditary colon cancer. *Science* 263, 1625-1629

Pearson, C.E. and Sinden, R.R. (1996). Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile X loci. *Biochemistry* 35, 5041-5053

Pogribny, I.P., Basnakian, A.G., Miller, B.J., Lopatina, N.G., Poirier, L.A., and James, S.J. (1995). Breaks in genomic DNA and within the p53 gene are associated with hypomethylation in livers of folate/methyl-deficient rats. *Canc Res* 55, 1894-1901

Pulst, S.M., Nechiporuk, A., Nechiporuk, T., Gispert, S., Chen, X.N., Lopescendes, I., Pearlman, S., Starkman, S., Orozcodiaz, G., Lunkes, A., Dejong, P., Rouleau, G.A., Auburger, G., Korenberg J.R., Figueroa, C., and Sahba, S. (1996). Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat Genet* 14, 269-276

Reddy, S., Smith, D.B., Rich, M.M., Leferovich, J.M., Reilly, P., Davies, B.M., Tran, K., Rayburn, H., Bronson, R., Cros, D., Balice-Gordon R.J., and Housman, D. (1996). Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. *Nat Genet* 13, 325-335

Redman, J.B., Fenwick, R.G., Fu, Y.-H., Pizzuti, A., and Caskey, T. (1993). Relationship between parental trinucleotide CTG repeat length and severity of myotonic dystrophy in offspring. *JAMA* 269, 1960-1965

Reiss, A.L., Kazazian, H.H., Krebs, C.M., McAughan, A., Boehm, C.D., Abrams, M.T., and Nelson, D.L. (1994) Frequency and stability of the fragile X premutation. *Hum Mol Genet* 3, 393-398

Reyniers, E., Vits, L., De Boulle, K., Van Roy, B., Van Venzen, D., de Graaff, E., Verkerk, A.J., Jorens, H.Z., Darby, J.K., and Oostra, B. (1993). The full mutation in the FMRl gene of male fragile X patients is absent in their sperm. *Nat Genet* 4, 143-146

Richards, R.I. and Sutherland, G.R. (1992). Dynamic mutations: a new class of mutations causing human disease. *Cell* 70, 709-712

Richards, R,I. and Sutherland G.R. (1994). Simple repeat DNA is not replicated simply. *Nat Genet* 6, 114-116

Rousseau, F., Heitz, D., Oberlé, I., and Mandel, J.L. (1991). Selection in blood cells from female carriers of the fragile X syndrome: reverse correlation between age and proportion of active X chromosomes carrying the full mutation. *J Med Genet* 28, 830-836

Rousseau, F., Heitz, D., Tarleton, J., MacPherson, J., Malmgren, H., Dahl, N., Barnicoat, A., Matheu, C., Mornet, E., Tejada, I., Maddalena, A., Spiegel, R., Schinzel, A., Marcos, J.A.G., Schorderet, D.F., Schaap, T., Maccioni, L., Russo, S., Jacobs, P.A., Schwartz, P.A., Mandel, J.L., and Sherman, S. (1994). Higher rate of transition from fragile X premutations into full mutations in males than in females suggest post conceptional expansion of the CGG repeats. *Am J Hum Genet* Suppl 55:A240

Rousseau, F., Rouillard, P., Morel, M.-L., Khandjian, E.W., and Morgan, K. (1995). Prevalence of carriers of premutation-size alleles of the *FMR1* gene-and implications for the population genetics of the fragile X syndrome. *Am J Hum Genet* 57, 1006-1018

Royle, N.J., Clarkson, R., Wong, Z., and Jeffreys, A.J. (1987). Preferential localization of hypervariable minisatellites near human telomeres. *Cytogenet Cell Genet* 46, 685

Sabourin, L.A., Mahadavan, M.S., Narang, M., Lee, D.S.C., Surh, L.C., and Korneluk, R.G. (1993). Effect of the myotonic dystrophy (DM) mutation on mRNA levels of the DM gene. *Nature Genet* 4, 233-238

Sabourin, L.A. (1996). Myotonic Dystrophy: a study of the expression of the myotonic dystrophy gene in affected tissues and cells. PhD Thesis, University of Ottawa.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular cloning: A Laboratory Manual* (2nd Ed.). Cold Spring Harbor, NY

Sanpei, K., Takano, H., Igarashi, S., Sato, T., Oyake, M., Sasaki, H., Wakisaka, A., Tashiro, K., Ishida, Y., Ikeuchi, T., Koide, R., Saito, M.m Sato, A., Tanaka, T., Hanyu, S., Takiyama, Y., Nishizawa, M., Shimizu, N., Nomura, Y., Segawa, M., Iwabuchi, K., Eguchi, I., Tanaka, H., Takahashi, H., and Tsuki, S. (1996). Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and direct cloning technique. *Nat Genet* 14, 277-284

Schonk, D., Corwinkel-Dreiessen, M., van Dalan, I., Oerlemans, F., Smeets, B., Scepens, J., Huselbos, T., Cockburn, D., Boyd, Y., Davis, M., Retting, W., Shaw, D., Roses, A., Ropers, H., and Wieringa, B. (1989). Definition of subchromosomal intervals around the myotonic dystrophy region at region 19q. *Genomics* 4, 384-396

Shaw, D.J., Chaudhary, S., Rundle, S.A., Crow, S., Brook, J.D., Harper, P.S., and Harley H.G. (1993). A study of DNA methylation in myotonic dystrophy. *J Med Genet* 30, 189-192

Shelbourne, P., Davies, J., Buxton, J., Anvret, M., Blennow, E., Bonduelle, M., Schmedding, E., Glass, I., Lindenbaum, R., Lane, R., Williamson, R., and Johnson, K. (1993). Direct diagnosis of myotonic dystrophy with a disease-specific DNA marker. *The New England J of Med* 328, 471-475

Shutler, G.G., Korneluk, R.G., Tsilfidis, C., Mahadevan, M., Bailly, J., Smeets, H., Jansen, G., Wieringa, B., Lohman, F., Aslanidis, C., and de Jong, P.J. (1992). Physical mapping and cloning of the proximal segment of the myotonic dystrophy gene region. *Genomics* 13, 518-525

Shutler, G.G. (1993). Genetic and physical mapping of the myotonic dystrophy locus on human chromosome 19q13.3. PhD Thesis, University of Ottawa.

Singhal, R.K., Prasad, R., and Wilson, S.H. (1995). DNA polymerase β conducts the gap-filling step in uracil initiated base excision repair in a bovine testis nuclear extract. *J Biol Chem* 270, 949-959

Smeets, H.J.M., Brunner, H.G., Ropers, H.H., and Wieringa, B. (1989). Use of variable sequence motifs as genetic markers: application to study of myotonic dystrophy. *Hum Genet 83*, 245-251

Smit, M., van der Kooij-Meijs, E., Frants, R.R., Havekes, L., and Klasen, E. (1988). Apolipoprotein gene cluster on chromosome 19. Definite localization of the ApoC2 gene and the polymorphic HpaI site associated with type II hyperlipoproteinemia. *Hum Genet* 78, 90-93

Snow, K., Doud, L.K., Hagerman, R., Pergolizzi, R.G., Erster, S.H., and Thibodeau, S.N. (1993). Analysis of a CGG sequence at the FMR1 locus in fragile X families and the general population. *Am J Hum Genet* 53, 1217-1228

Spiegel, R., La Spada, A.R., Kress, W., Fischbeck, K.H., and Schmid, W. (1996). Somatic stability of the expanded CAG trinucleotide repeat in X-linked spinal and bulbar muscular atrophy. *Hum Mutat* 8, 32-27

Stallings, R.L., Ford, A.F., Nelson, D., Torney, D.C., Hildebrand, C.E., and Moyzis, R.K. (1991). Evolution and distribution of (GT)n repetitive sequences in mammalian genomes. *Genomics* 10, 807-815

Stark, G.R. (1993). Regulation and mechanisms of mammalian gene amplification. *Advances in Cancer Research* 61, 87-113

Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E., and Inouye, M. (1966). Frameshift mutations and the genetic code. *Cold Spring Harb. Symp. Quant. Biology* 31, 77-84

Sutherland, G.R. and Hecht, F. (1985). *Fragile sites on human chromosomes*, Oxford, Oxford Univ. Press.

Tachi, N. (1996). DNA diagnosis in myotonic dystrophy. *Hokkaido Igaku Zasshi* 71, 3-8

Taneja, K.L., McCurrach, M., Schalling, M., Housman, D., and Singer, R.H. (1995). Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol* 128, 995-1002

Thornton, C.A., Johnson, K., and Moxley, R.T. (1994a). Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. *Annals of Neurol* 35, 104-107

Thornton, C.A., Griggs, R.C., and Moxley, R.T. (1994b). Myotonic dystrophy with no trinucleotide repeat expansion. *Annals of Neurol* 35, 269-272

Tsilfidis, C., MacKenzie, A.E., Mettler, G., **Barceló, J.**, and Korneluk, R.G. (1992). Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. *Nature Genet* 1, 192-195

van den Ouweland, A.M., Deelen, W.H., Kunst, C.B., Uzielli, M.L., Nelson, D.L., Warren, S.T., Oostra, B.A., and Halley, D.J. (1994). Loss of mutation at the FMR1 locus through multiple exchanges between maternal X chromosomes. *Hum Molec Genet* 3, 1823-1827

van der Ven, P.F.M., Jansen, G., van Kuppevelt, T.H.M.S.M., Perryman, M.B., Lupa, M., Dunne, P.W., ter Laak, H.J., Jap, P.H.K., Veerkamp, J.H., Epstein, H.J., and Wieringa, B. (1993). Myotonic dystrophy kinase is a component of neuromuscular junctions. *Hum Molec Genet* 2, 1889-1894

Vanier, T.M. (1960). Dystrophia Myotonica in Childhood. *Br Med J* 2, 1284-1288

Verkerk, A.J.M.H., Pieretti, M., Sutcliffe, J.S., Fu, Y-H., Kuhl, D.P.A., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F., Eussen, B.E., van Ommen, G-J.B., Blonden, L.A.J., Riggins, G.J., Chastain, J.L., Kunst, C.B., Galjaard, H., Caskey, C.T., Nelson, D.L., Oostra, B.A., and Warren, S.T. (1991). Identification of a gene (*FMR-1*) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in Fragile-X syndrome. *Cell* 65, 905-914

Waga, S., Bauer, G., and Stillman, B. (1994). Reconstitution of compiete SV40 DNA replication with purified replication factors. *J Biol Chem* 269, 10923-10934

Wang, J., Pegoraro, E., Menegazzo, E., Genarelli, M., Hoop, R.C., Angineli, C., and Hoffman, E.P. (1995). Myotonic dystrophy: evidence for a possible dominant-negative RNA mutation. *Hum Molec Genet* 4, 599-606

Wang, Y.-H. and Griffith, J. (1995). Expanded CTG triplet blocks from the myotonic dystrophy gene create the strongest known natural nucleosome positioning elements. *Genomics* 25, 570-573

Wang, Y.-H., Amirhaeri, S., Kang, S., Wells, R.D., and Griffith, J.D. (1994b). Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science* 265, 669-671

Waring, J.D., Hag, R., Tamai, K., Sabourin, L.A., Ikeda, J.-E., and Korneluk, R.G. (1996). Investigation of myotonic dystrophy kinase isoform translocation and membrane association. *The J of Biol Chem* 271, 15187-15193

Warren, S.T. (1996). The expanding world of trinucleotide repeats. *Science* 271, 1374-1375

Weber J.L. and May, P.E. (1988). An abundant new class of human DNA polymorphisms. *Am J Hum Genet* (Suppl) 43:A161

Wells, R.D., Ohtsuka, E., and Khorana, H.G. (1965). Studies on polynucleotides. *J Molec Biol* 14, 221-240

Weng, Y.S., Whelden, J., Gunn, L., and Nickoloff, J.A. (1996). Double-strand break-induced mitotic gene conversion: examination of tract polarity and products of multiple recombinational repair events. *Curr Genet* 29, 335-343

Whiting, E.J., Waring, J.D., Tamai, K., Somerville, M.J., Hincke, M., Staines, W.A., Ikeda, J.-E., and Korneluk, R.G. (1995a). Characterization of myotonic dystrophy kinase (DMK) protein in human and rodent muscle and central nervous tissue. *Hum Molec Genet* 4, 1063-1072

Whiting, E.J., Tsifidis, C., Surh, L., Mackenzie, A.E., and Korneluk, R.G. (1995b). Convergent myotonic dystrophy (DM) haplotypes: potential inconsistencies in human disease gene localization. *Europ Jour Hum Genet* 3, 195-202

Wieringa, B (1994). Myotonic dystrophy reviewed: back to the future? *Hum Molec Genet* 3, 1-7

Wöhrle, D., Hennig, I., Vogel, W., and Steinbach, P. (1993). Mitotic stability of fragile X mutations in differentiated cells indicates early post-conceptional trinucleotide repeat expansion. *Nat Genet* 4, 140-142

Wöhrle, D., Kennerknecht, I., Wolf, M., Enders, H., Schwemmle, S., and Steinbach, P. (1995). Heterogeneity of DM kinase repeat expansion in different fetal tissues and further expansion during cell proliferation in vitro: evidence for a causal involvement of methyl-directed DNA mismatch repair in triplet repeat stability. *Hum Mol Genet* 4, 1147-1153

Wolff, R., Nakamura, Y., and White, R. (1988). Molecular characterization of a spontaneously generated new allele at a VNTR locus: no exchange of flanking DNA sequence. *Genomics* 3, 347-351

Wolff, R.K., Plaetke, R., Jeffreys, A.J., and White, R. (1989). Unequal crossing-over between homologous chromosomes is not the major mechanism involved in the generation of new alleles at the VNTR loci. *Genomics* 5, 382-384

Wong, L.-J. C., Ashizawa, T., Monckton, D.G., Caskey, C.T., and Richards, C.S. (1995). Somatic heterogeneity of the CTG repeat in myotonic dystrophy is age and size dependent. *Am J Hum Genet* 56, 114-122

Yano-Yanagisawa, H., Li, Y., Wang, H., and Kohwi, Y. (1995). Single-stranded DNA binding proteins isolated from mouse brain recognize specific trinucleotide repeat sequences in vitro. *Nucleic Acids Res* 23, 2654-2660

Yu, S., Mulley, J., Loesch, D., Turner, G., Donnelly, A., Gedeon, A., Hillen. D., Kremer, E., Lynch, M., Pritchard, M., Sutherland, G.R., and Richards, R.I. (1991). Fragile-X syndrome: unique genetics of the heritable unstable element. *Am J Hum Genet* 50, 968-980.

Zatz, M., Passos-Bueno, M.R., Cerqueira, A., Marie, S.K., Vainzof, M., and Pavanello, R.C. (1995). Analysis of the CTG repeat in skeletal muscle of young and adult myotonic dystrophy patients: when does the expansion occur? *Hum Mol Genet* 4, 401-406

Zuliani, G. and Hobbs, H.H. (1990). A high frequency of length polymorphisms in repeated sequences adjacent to Alu sequences. *Am J Hum Genet* 46, 963-969

APPENDIX: IMPROVEMENT OF PCR YIELDS IN THE MYOTONIC DYSTROPHY AMPLIFIED CTG REPEAT SEQUENCE BY A SINGLE-STRANDED DNA BINDING PROTEIN

1. Introduction

The molecular diagnosis of DM requires the estimation of the size of the mutant allele, which is determined by the number of trinucleotide repeats. Since the size of the mutation correlates positively with the severity of the clinical symptoms in DM, its estimation is important. Two complementary procedures are used for the molecular diagnosis of myotonic dystrophy (Surh et al., 1994): the first consists of the Southern transfer of restriction enzyme-digested DNA to nylon membranes followed by detection of the specific DMK band containing the (CTG)n repeat tract with a labelled probe. The second procedure consists of PCR amplification of the repeat sequence with flanking primers followed by Southern transfer to a nylon membrane and detection with a labelled (CTG)n or (CAG)n oligonucleotide probe. Both procedures have been shown to be necessary in most cases (Surh et al., 1994). The first method is needed for the detection of large mutant alleles and the second one gives a much more accurate measurement of small mutant alleles which are otherwise often indistinguishable from normal alleles in genomic Southern blot analysis. A faster, simpler, and less costly mutation detection protocol, based exclusively on the non-isotopic detection of PCR products directly in an agarose gel, is currently not feasible due to the poor visibility of the mutant allele bands. Standard PCR protocols applied to the detection of large DM alleles result in very faint or invisible smears on agarose gels. These smears are clearly

visible only after Southern blot transfer of this DNA to nitrocellulose membranes followed by specific probe-mediated detection (see Mahadevan et al., 1992, for examples).

The following describes a protocol in which gene 32 protein (gp32, a single-stranded DNA binding protein from bacteriophage T4) is added to the PCR reaction mixture. This results in a 2- to 25-fold increase in the amount of product obtained from the PCR amplification of this CTG repeat sequence. In contrast, when the same protocol is applied to DNA templates containing non-repeat sequences, the yield of the PCR products decreases.

MATERIALS AND METHODS

Preparation of Genomic DNA

Genomic DNA was obtained from peripheral blood leukocytes following a standard phenol-chloroform extraction method and diluted in TE to a concentration of approximately 500 ng/µl.

Reaction Mixture

A 50 µl PCR reaction mixture contained 1.5 µl of genomic DNA, 22 µl of distilled H_2O, 5 µl of 10 x Taq polymerase buffer with 10 mM Tris-HCl pH 8.3, 50 mM KCl, 15 mM $MgCl_2$, 5 µl of each of either one of the oligonucleotide primer sets 406 and 409 or 409 and 410 (at an initial concentration of 20 µg/µl) (Mahadevan et al., 1992),

8 μl of a dNTP mix containing the four precursors at a concentration of 1.25 mM each, 2.5 μl of DMSO, 0.5 μl of Taq polymerase (Cetus, 5 U/μl) and 0.5 μl of gp32 (Pharmacia Biotech, 4.1 μg/μl).

Amplification Conditions

A Perkin-Elmer Cetus 9600 Gene-Amp thermocycler was used for the experiments. The reaction mixture was submitted to a program that included 3 min of DNA denaturation at 95°C followed by 25 cycles of 30 s at 95°C, 20 s at 61°C and 30 s at 72 °C followed by 10 min extension at 72°C. The oligonucleotide primers employed in the experiment designed to check on the specificity of gp32 for the CTG repeat, and described in Figure A-2, are the following: Set A: 458 (5'-CTGCAGAAGGTTTAGAAAGAGC-3') and 424 (5'-TCATCCTGTGGGACACCGAGG-3'), set B: 414 (5'-GTTCGCAAAGTGCAAAGCTTTCT-3') and 417 (5'-TTCCCGGCTACAAGGACCCTTC-3'), set C: 485 (5'-TGCCACCCGCTTAGCTGCG-3') and 417 (sequences from the DMK gene, GenBank accession number L08835).

Detection and Quantification of Products

A total of 10 μl of the PCR products were submitted to electrophoresis on a 1.5% agarose gel containing ethidium bromide. Quantitative measurement of the relative efficiency of the PCR reaction was performed by densitometry scanning (with BIO-RAD Model GS-670 Imaging Densitometer) of the polaroid image of the agarose gel containing the PCR products.

3. Results and discussion

This simple modification of the basic PCR protocol allows, and in fact favours, the use of PCR primers that flank very closely the CTG repeat sequence. The advantage of using such primers is that the number of repeats can be directly and precisely determined without having to resort to restriction digestions and/or Southern transfer followed by probe-mediated detection.

An example of the effect of the presence of gp32 in the reaction is shown in Figure A-1. The increase in the amount of product in the mutant allele band varied from 2 to 25 fold. The protocol allows for direct detection and sizing of bands in the gels. This is in contrast to PCR protocols applied to templates with high G/C content in which inclusion of the nucleotide analogue 7'-deaza-2'-deoxyguanosine-5'-triphosphate (C^7dGTP) is used as a strategy to allow for a more efficient amplification of the expected sequence. With these protocols additional steps are often needed in order to see the obtained products since DNA containing this nucleotide analogue have a very low capacity to emit fluorescent radiation (Weiss et al., 1994).

The increased efficiency of the PCR reaction with this repeat sequence in the presence of gp32 is not due to an effect of this product on the oligonucleotide primers. Different primer sets flanking the CTG repeat yielded comparable amounts of PCR product. A striking reduction in PCR efficiency with gp32 was found when randomly selected DNA templates that did not contain the same motif (Figure A-2). These results

Figure A-1. Analysis of the effect of gp32 on the efficiency of PCR amplification on the CTG repeat template: a) PCR reaction conditions as described using oligonucleotide primers 406 and 409 (Mahadevan et al., 1992). A 10 µl aliquot of each product was analysed by electrophoresis on a 1% agarose gel which was subsequently stained with ethidium bromide. **A, B, C** and **D** correspond to the DNA of four different DM patients. Upper and lower bands correspond to the PCR products of the mutant (**m**) and normal (**n**), alleles respectively, of the same individuals. The mutant alleles contained approximately 170 (**A**), 250 (**B**), 76 (**C**) and 235 (**D**) CTG repeats and normal alleles contained 10 (**A**), 5 (**B**), 11 (**C**) and 24 (**D**) CTG repeats. Reaction products in which gp32 was added prior to the PCR amplification are shown in lanes **2**, **4**, **6** and **8**. Lanes **1**, **3**, **5** and **7** correspond to the same reactions without gp32. The increase in the yield of the mutant allele band in PCR reactions containing gp32 ranged from 2 to 25 fold. With larger-size normal alleles the increase in efficiency of the PCR reaction is also evident with gp32. **b)** Autoradiograph of products described in **a)** transferred to a nitrocellulose membrane and probed with an alkaline phosphatase conjugated $(CAG)_{10}$ oligonucleotide. The smeared appearance of the upper bands is a characteristic of this PCR assay on DNA originating from DM patients and is due to both somatic heterogeneity and the technical difficulty in amplification of these CTG repeats.

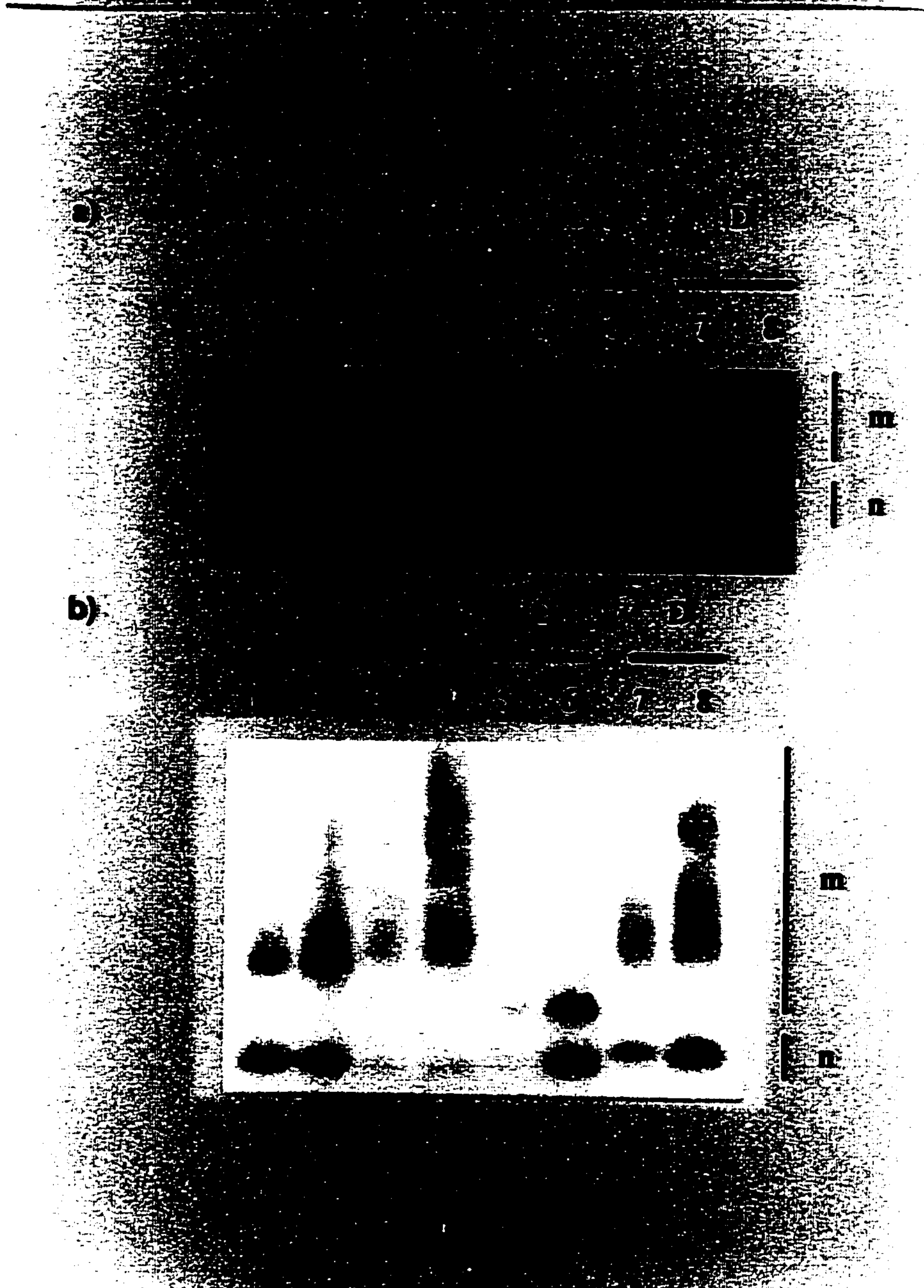
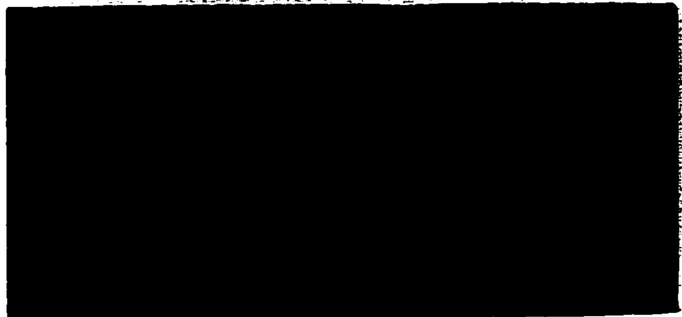

a)
m
n
b)
m
n

Figure A-2. Analysis of the specificity of the effect of gp32 on different DNA templates. Conditions were as described as for Figure A-1 except that three different sets (**A**, **B** and **C**) of oligonucleotide primers were chosen in a PCR reaction so that the DM CTG repeat would not be part of the template. The decrease in the yield of PCR products is obvious in reaction mixtures containing gp32 (lanes **2**, **4** and **6**). The numbers on either side of the photograph represent the size of the bands in base-pairs. The differences between the results of Figure A-1 and this figure reflect the specificity of the interaction of gp32 with the CTG trinucleotide repeat.


A
B
C
1
2
3
4
5
6
908
350
647

suggest a specific interaction of gp32 with the DM CTG repeat sequence that somehow affect its amplification. Long stretches of CTG repeats may result in an unusual DNA conformation which makes their PCR amplification inefficient. The higher yields of PCR products containing this sequence, in the presence of gp32, may indicate the partial disruption of this structure.

The difficulty in cloning large stretches of CTG repeats into plasmids is partly due to the difficulty of obtaining enough repeat-containing PCR product. In my experience the use of gp32 in the PCR amplification of patient material improves significantly the cloning efficiency of relatively large CTG repeat sequences.

The structural basis of the interaction between gp32 and the CTG repeats remains to be determined. There is considerable interest in the structure adopted by this DNA sequence and study of its interaction with gp32 and other nucleic-acid-binding proteins may shed light on this problem. Ultimately, knowledge of the structure adopted by large stretches of CTG repeats *in vivo* may be a key to the understanding of the mechanism of instability of this sequence leading to the disease.

Although the described method cannot entirely substitute for the one(s) in current use for the molecular diagnosis of DM (Surh et al., 1994), we found that the addition of gp32 considerably improves the results obtained with the PCR portion of the current mutation detection protocols. As a result, this modification of the PCR protocol is being

used on a regular basis for the diagnosis of DM at the Molecular Genetics Diagnostics laboratory at the Children's Hospital of Eastern Ontario in Ottawa.

REFERENCES

Mahadevan M.S., C. Tsilfidis, L. Sabourin, G. Shutler, C. Anemiya, G. Jansen, C. Neville, M. Narang, **J. Barceló**, K. O'Hoy, S. Leblond, J. Earle-MacDonald, P.J. de Jong, B. Wieringa, and R.G. Korneluk. (1992). Myotonic dystrophy mutation: and unstable CTG repeat in the 3' untranslated region of the gene. *Science* 255, 1253-1255

Surh, L.C., M. Mahadevan, and R.G. Korneluk. (1994). Analysis of trinucleotide repeats in myotonic dystrophy, p. 9.6.1-9.6.8. *In* N.C. Dracopoli et al (Eds), Current Protocols in Human Genetics, Vol. 1. John Willey & Sons, Inc. USA.

Weiss, J., H.-D. Zucht, and W.-G. Forssmann. (1994). Amplification of gene fragments with very high G/C content: c^7dGTP and the problem of visualizing the amplification products. *PCR Methods and Applications* 4, 124-12502