

**University of Glasgow  
Institute of Biomedical and Life Sciences  
Division of Molecular Genetics**

**Genetic and Environmental Modifiers of  
Somatic Trinucleotide Repeat Dynamics**

**Mário Gomes-Pereira**

**Thesis submitted for the degree of Doctor of Philosophy**

**June 2002**

ProQuest Number: 13833915

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

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



ProQuest 13833915

Published by ProQuest LLC (2019). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

**GLASGOW  
UNIVERSITY  
LIBRARY:**

# Abstract

The expansion of CAG•CTG trinucleotide repeat sequences has been identified as the genetic cause of several human diseases, including myotonic dystrophy type 1, Huntington disease, and an ever-increasing number of spinocerebellar ataxias. Once above a size threshold, the repeats become dramatically unstable in the germline and also throughout the soma, with a marked bias towards further expansion. Such expansions constitute a unique form of dynamic mutation, whose mechanism is poorly understood. While germline instability serves as the molecular basis for genetic anticipation; age-dependent, tissue-specific somatic instability most likely contributes to the tissue specificity, phenotypic variability and progressive nature of these conditions. The study of the mutation mechanism is therefore of major interest, as it may provide valuable clues towards a better understanding of disease pathophysiology.

It is generally assumed that the repeat length changes arise through DNA polymerase slippage during DNA replication, however no direct evidence exists to support this hypothesis in mammalian systems. Transgenic mouse models of unstable CAG•CTG repeats have been previously generated, and shown to recreate the dynamic nature of somatic mosaicism observed in humans. Tissues from these mice have now been used in order to establish an *in vitro* cell culture system, where the repeat dynamics could be investigated under controlled conditions. Monitoring of repeat stability in these cells over long periods of time, and numerous population doublings, has revealed the progressive accumulation of larger alleles, as a result of repeat length changes *in vitro*, confirmed by single cell cloning. Selection of cells carrying longer repeats was observed during the first few passages of the cultures, and frequent additional selective sweeps were also detected at later stages. The highest levels of instability were observed in cultured kidney cells, whilst the transgene remained relatively stable in eye cells and very stable in lung cells, paralleling previous *in vivo* observations. More importantly, the levels of repeat instability in cultured cells did not correlate with cell proliferation rates, rejecting a simple association between length change mutations and cell division, and suggesting an important role for additional cell type-specific and possibly environmental *trans*-acting modifiers of repeat metabolism.

The effects of multiple genotoxic agents on the mutational dynamics of expanded trinucleotide repeats were assessed in this tissue culture model of unstable DNA. The drugs tested were selected based on their ability to affect cell cycle progression, DNA polymerase activity, DNA methylation, intracellular levels of oxidative stress or DNA conformational metabolism. The analysis led to the identification of chemicals, such as aspirin, 5-azacytidine and 1-β-D-arabinofuranosyl-cytosine that resulted in the deceleration of the rate of trinucleotide repeat expansion, particularly in a kidney clonal cell line carrying rapidly expanding repetitive tracts. These observations were reported in the absence of major changes in the rates of cell turnover. In contrast, forced cell cycle progression by exposure to caffeine resulted in a significantly higher rate of triplet repeat expansion. Increased levels of oxidative stress, generated in culture by exposure to

a variety of drugs, were associated with reduced levels of repeat size variability, most likely through means of cell selection in culture. Since pathology in CAG•CTG-associated diseases is mediated by a variety of complex and unrelated molecular pathways, drug induced modification of DNA dynamics could present a possible therapeutical route for these disorders. Specifically, chemical treatments that resulted in suppression of somatic repeat expansion would be expected to be beneficial, whilst reversion of the expanded mutant repeat to the normal repeat size range, observed in the general population, would be predicted to be curative. Although preliminary, the findings described in this study may open new avenues in the search for novel therapeutical strategies.

Mechanistic models of repeat length mutation based on DNA replication, recombination and repair have been proposed. The latter have implied the involvement of mammalian *MutS* homologues (*Msh2*, *Msh3* and *Msh6*). In order to gain further insight into the molecular mechanisms driving trinucleotide repeat mutation, the involvement of a mammalian *MutL* homologue (*Pms2*) in the mutation dynamics was investigated. No significant differences were observed between *Pms2*<sup>+/+</sup> and *Pms2*<sup>+/-</sup> mice, suggesting that a single functional *Pms2* allele is sufficient to maintain high levels of somatic mosaicism. The levels of *Pms2* mRNA and protein in heterozygotes deficient for *Pms2* have not yet been investigated. In contrast to what would be predicted by the replication slippage model, lower levels of trinucleotide somatic mosaicism were detected in homozygous *Pms2*-null mice, compared with age-matched controls, carrying either one or two functional copies of the *Pms2* allele. In addition, a higher frequency of rare but large deletion events was detected in *Pms2*<sup>-/-</sup> animals. Both results proved statistically significant by single molecule analysis. These findings imply that, not only MMR enzymes that directly bind to DNA, but also proteins that are subsequently recruited by MutS proteins, play a central role in the accumulation of repeat length changes, arguing against a mutation mechanism mediated by stabilisation of alternative DNA secondary structures by MMR proteins. MMR gene polymorphisms and variants might therefore be considered potential determinants of trinucleotide repeat instability in humans, predicted to affect both age of onset and disease progression.

*Let us learn to dream.... then perhaps we shall find the truth.*

Friedrich August Kekulé

*Também eu, também eu,  
joguei às escondidas, fiz baloiços,  
tive bolas, berlindes, papagaios,  
automóveis de corda, cavalinhos...*

*Depois cresci,  
tornei-me do tamanho que hoje tenho,  
os brinquedos perdi-os, os meus bibes  
deixaram de servir-me.  
Mas nem tudo se foi:  
ficou-me,  
dos tempos de menino,  
esta alegria ingénua  
perante as coisas novas  
e esta vontade de brincar.*

*Vida!  
não me venhas roubar o meu tesouro:  
não te importes que eu ria,  
que eu salte como dantes.  
E se eu riscar os muros  
ou quebrar algum vidro  
ralha, ralha comigo, mas de manso...*

*Eu tinha um bibe azul...  
Tinha berlindes,  
tinha bolas, cavalos, papagaios...  
A minha Mãe ralhava assim como quem beija...  
E quantas vezes eu, só pra ouvi-la  
ralhar, parti os vidros da janela  
e desenhei bonecos na parede...*

*Vida!, ralha também,  
ralha, se eu te fizer maldades, mas de manso,  
como se fosse ainda a minha Mãe...*

Sebastião da Gama

# Table of contents

<b>Abstract</b> .....	<b>2</b>
<b>Table of contents</b> .....	<b>5</b>
<b>List of tables</b> .....	<b>10</b>
<b>List of figures</b> .....	<b>11</b>
<b>Acknowledgements</b> .....	<b>14</b>
<b>List of Abbreviations</b> .....	<b>18</b>
<b>1. Introduction</b> .....	<b>21</b>
1.1. <i>Trinucleotide repeats and human disease</i> .....	21
1.2. <i>Molecular pathogenesis</i> .....	24
1.2.1. Polyglutamine diseases .....	24
1.2.1.1. Toxicity of polyglutamine tracts.....	25
1.2.1.2. Polyglutamine protein intranuclear inclusions.....	26
1.2.1.3. Cytoplasmic aggregates .....	26
1.2.1.4. Cell dysfunction versus cell death.....	27
1.2.1.5. Cell-specific vulnerability in polyglutamine diseases .....	27
1.2.2. Non-coding repeat disorders.....	28
1.2.2.1. Fragile X syndrome (FRAXA) .....	28
1.2.2.2. Fragile XE mental retardation (FRAXE) .....	29
1.2.2.3. Friedreich ataxia (FRDA) .....	29
1.2.2.4. Myotonic dystrophy type 1 (DM1).....	30
1.2.2.5. Spinocerebellar ataxia type 8 (SCA8).....	34
1.2.2.6. Spinocerebellar ataxia type 12 (SCA12).....	35
1.3. <i>Repeat dynamics</i> .....	35
1.3.1. Germline instability .....	35
1.3.2. Somatic instability .....	37
1.3.3. <i>Dmt</i> transgenic mice .....	39
1.3.4. Possible link between somatic instability and pathophysiology .....	39
1.4. <i>Modifiers of trinucleotide repeat instability</i> .....	41
1.4.1. <i>Cis</i> -acting modifiers of trinucleotide repeat instability .....	41
1.4.2. <i>Trans</i> -acting modifiers of trinucleotide repeat dynamics.....	43
1.5. <i>DNA repair systems</i> .....	44
1.5.1. Base excision repair .....	44
1.5.2. Nucleotide excision repair .....	44
1.5.3. DNA mismatch repair .....	45
1.5.3.1. MMR, microsatellite instability and cancer .....	47
1.5.3.2. MMR proteins in cell cycle surveillance and apoptosis .....	49
1.6. <i>Mechanistic models of trinucleotide repeat expansion</i> .....	49
1.6.1. DNA polymerase slippage model.....	50
1.6.1.1. Transcription interference in repeat length changes .....	52
1.6.1.2. FEN1 interference model .....	53
1.6.2. Recombination-dependent mechanisms of expansion.....	55
1.6.2.1. Generation of DNA strand breaks within trinucleotide repeat sequences .....	56
1.6.2.2. Trinucleotide repeat instability by double strand break repair.....	56

1.6.3. Models of germline trinucleotide repeat instability .....	58
1.7. <i>Project design</i> .....	60
<b>2. Materials and methods</b> .....	<b>62</b>
2.1. <i>General materials</i> .....	62
2.1.1. Chemicals and reagents .....	62
2.1.2. Immunochemicals .....	62
2.1.3. General disposable plasticware .....	63
2.1.4. Tissue culture disposable materials .....	63
2.1.5. Enzymes .....	64
2.1.6. Nucleic acids size markers and mass ladders .....	64
2.1.7. Kits.....	65
2.1.8. Membranes and paper .....	65
2.1.9. Photography and autoradiography.....	65
2.1.10. Microscopes.....	65
2.1.11. Spectrophotometer.....	66
2.1.12. DNA crosslinker.....	66
2.2. <i>Experimental materials</i> .....	66
2.2.1. Animals .....	66
2.2.2. Bacterial host strains.....	66
2.2.3. Vectors.....	66
2.2.4. DNA sources .....	66
2.2.5. Oligonucleotides .....	67
2.2.6. Probes .....	68
2.3. <i>Solutions</i> .....	68
2.3.1. General solutions.....	68
2.3.2. Bacterial solutions, media and antibiotics.....	71
2.3.3. Tissue culture solutions, media and antibiotics .....	72
2.4. <i>Mouse tissue culture methods</i> .....	72
2.4.1. Establishment of primary cell cultures.....	72
2.4.1.1. Eye cell cultures .....	73
2.4.1.2. Lung and kidney cell cultures.....	73
2.4.2. Establishment of single cell-derived clones.....	73
2.4.3. Feeding cultured cells .....	74
2.4.4. Subculturing cultured cells .....	74
2.4.5. Measuring cell counts and determining population doubling times .....	74
2.4.6. Measuring cell viability .....	75
2.4.7. Freezing cultured mouse cells in liquid nitrogen.....	75
2.4.8. Thawing cultured mouse cells .....	76
2.4.9. Immunocytochemical characterisation of cultured cell types .....	76
2.4.10. 5'-Bromo-3-deoxyuridine (BrdU) incorporation and detection assay .....	76
2.5. <i>Preparation, purification and analysis of DNA</i> .....	77
2.5.1. Preparation of plasmid DNA .....	77
2.5.2. DNA extraction from mouse tissues .....	77
2.5.3. DNA extraction from cultured mouse cells .....	78
2.5.4. Determination of DNA concentration.....	78
2.5.5. Polymerase chain reaction (PCR).....	79
2.5.6. DNA cloning techniques.....	79
2.5.6.1. Restriction endonuclease digestion of plasmid DNA .....	79
2.5.6.2. DNA ligation .....	79
2.5.6.3. Transformation of competent bacterial cells.....	80
2.5.6.4. Generation of plasmid stocks.....	80
2.5.7. Gel electrophoresis.....	80
2.5.7.1. Agarose gel electrophoresis .....	80
2.5.7.2. Non-denaturing polyacrylamide gel electrophoresis .....	80
2.5.7.3. DNA extraction from non-denaturing polyacrylamide gels .....	81
2.5.8. Southern blotting .....	81

2.5.8.1. Genomic DNA transfer from agarose gels onto a nylon membrane by Southern blotting .....	81
2.5.8.2. DNA transfer from agarose or polyacrylamide gels onto a nylon membrane by Southern “squash” blotting .....	82
2.5.9. Preparation of radiolabelled double-stranded probes .....	82
2.5.10. Southern hybridisation .....	83
2.5.10.1. Hybridisation of genomic DNA transferred onto a nylon membrane .....	83
2.5.10.2. Hybridisation of PCR products and plasmid DNA transferred onto a nylon membrane .....	83
2.6. <i>Preparation, purification and analysis of RNA</i> .....	84
2.6.1. RNA extraction from cultured mouse cells .....	84
2.6.2. Determination of RNA concentration .....	84
2.6.3. RNA visualisation by agarose gel electrophoresis .....	85
2.6.4. Northern blotting and hybridisation .....	85
2.6.4.1. RNA electrophoresis in denaturing agarose gels .....	85
2.6.4.2. RNA transfer from denaturing agarose gels onto a nylon membrane by northern blotting .....	85
2.6.4.3. Hybridisation of RNA transferred onto a nylon membrane .....	85
2.7. <i>Small pool PCR analysis</i> .....	86
2.7.1. Restriction endonuclease digestion of mouse genomic DNA .....	86
2.7.2. Small pool PCR amplification .....	86
2.7.3. Gel electrophoresis of SP-PCR products .....	87
2.7.4. DNA transfer from the gel onto nylon membranes by Southern “squash” blot .....	87
2.7.5. Single molecule SP-PCR analysis .....	87
<b>3. Establishment and characterisation of a mouse cell culture model system of trinucleotide repeat instability</b> .....	<b>89</b>
3.1. <i>Introduction</i> .....	89
3.2. <i>Results</i> .....	91
3.2.1. Mouse genotyping for the <i>Dmt162</i> transgene .....	91
3.2.2. Growth dynamics of mouse cell cultures .....	93
3.2.3. Characterisation of cultured cell types .....	95
3.2.4. Tissue-specific trinucleotide instability and selection for longer alleles in cultured mouse cells .....	95
3.2.4.1. Repeat size variability in lung cell cultures (D2763L cell line) .....	95
3.2.4.2. Repeat size variability in eye cell cultures (D2763E cell line) .....	97
3.2.4.3. Repeat size variability in kidney cell cultures (D2763K cell line) .....	97
3.2.5. Age-of-donor effect on trinucleotide repeat instability observed in culture .....	99
3.2.6. Accumulation of mutations in single cell-derived clones from kidney cell cultures .....	103
3.2.7. Preferential accumulation of longer alleles in competition assays between clones carrying different sized repeats .....	105
3.2.8. Repeat size distributions in kidney and liver tissue samples <i>in vivo</i> .....	107
3.2.9. Trinucleotide repeat stability in <i>Dmt-E</i> kidney cells .....	109
3.3. <i>Discussion</i> .....	109
<b>4. Investigating the role of DNA topology as a mediator of trinucleotide repeat dynamics</b> .....	<b>115</b>
4.1. <i>Introduction</i> .....	115
4.2. <i>Results</i> .....	119
4.2.1. Detection of DNA alternative structures in CAG•CTG sequences generated by PCR amplification .....	119
4.2.1.1. Analysis of alternative CAG•CTG-containing PCR products by native polyacrylamide gel electrophoresis .....	119
4.2.1.2. Confirmation of S-DNA folding within CAG•CTG sequences .....	123
4.2.1.3. Generation of alternative structures by PCR amplification .....	125
4.2.2. Testing RecA activity on CAG•CTG repetitive sequences .....	126
4.2.2.1. RecA activity assay on CAG•CTG repetitive tracts .....	126

4.2.3.	Novobiocin as a genotoxic modifier of trinucleotide repeat instability: a possible role for topoisomerase II in triplet repeat metabolism.....	128
4.2.3.1.	D2763K cell line .....	129
4.2.3.2.	D2763Kc2 cell line .....	131
4.2.3.3.	D4132K cell line .....	133
4.3.	<i>Discussion</i> .....	135
<b>5.</b>	<b>Genotoxic effects of oxidative stress on trinucleotide repeat dynamics .....</b>	<b>141</b>
5.1.	<i>Introduction</i> .....	141
5.2.	<i>Results</i> .....	147
5.2.1.	Association between levels of somatic instability and sensitivity to oxidative stress in culture .....	147
5.2.2.	Induction of trinucleotide repeat size variability <i>in vitro</i> by hydrogen peroxide: a preliminary study .....	148
5.2.2.1.	D2763K kidney cell line .....	149
5.2.2.2.	D3111K kidney cell line .....	151
5.2.2.3.	D2763E eye cell line .....	151
5.2.3.	Quantitative analysis of the effect of hydrogen peroxide-induced oxidative stress on the dynamics of expanded CAG•CTG repeats.....	152
5.2.3.1.	D2763Kc2 kidney cell line .....	152
5.2.3.2.	D4132K kidney cell line .....	154
5.2.4.	Effect of hydrogen peroxide treatment on stable expanded CAG•CTG repeats .....	156
5.2.5.	Triplet repeat instability in mouse cells with impaired mitochondrial function.....	157
5.2.5.1.	Effects of ethidium bromide on trinucleotide repeat dynamics.....	158
5.2.5.2.	Investigating the biological bases of the effects of ethidium bromide on trinucleotide repeat dynamics.....	160
5.2.5.3.	Altered electrophoretic mobility of trinucleotide repeat PCR products in the presence of ethidium bromide .....	161
5.2.5.4.	Searching for alternative DNA structures <i>in vivo</i> .....	165
5.2.5.5.	Effects of rhodamine-6G on trinucleotide repeat dynamics.....	168
5.3.	<i>Discussion</i> .....	171
<b>6.</b>	<b>Investigating the effects of multiple genotoxic agents, affecting cell cycle progression, DNA replication and DNA repair, on the dynamics of expanded triplet repeats .....</b>	<b>178</b>
6.1.	<i>Introduction</i> .....	178
6.2.	<i>Results</i> .....	182
6.2.1.	Association between increased sensitivity to UV radiation and greater levels of trinucleotide repeat instability .....	182
6.2.2.	Effects of exposure to UV radiation on the dynamics of expanded trinucleotide repeats .....	183
6.2.3.	Inhibition of DNA damage checkpoint by caffeine induces greater levels of trinucleotide repeat instability .....	188
6.2.4.	Inhibition of DNA polymerase elongation step by araC is associated with lower levels of expansion-biased repeat instability.....	194
6.2.5.	Induction of DNA hypomethylation by 5-azacytidine and its consequences on the dynamics of CAG•CTG repeats .....	194
6.2.6.	The effects of aspirin exposure on the dynamics of expanded CAG•CTG repeats.....	196
6.3.	<i>Discussion</i> .....	197
<b>7.</b>	<b>Trinucleotide repeat dynamics in slowly proliferating and non-dividing <i>Dmt-D</i> cells.....</b>	<b>205</b>
7.1.	<i>Introduction</i> .....	205
7.2.	<i>Results</i> .....	209
7.2.1.	Trinucleotide repeat dynamics in <i>Dmt-D</i> cells growing at high density in low foetal bovine serum .....	209

7.2.2. Trinucleotide repeat dynamics in confluent <i>Dmt-D</i> cell cultures arrested by contact inhibition .....	211
7.2.3. Trinucleotide repeat dynamics in <i>Dmt-D</i> cells arrested by mitomycin C.....	214
7.2.4. Trinucleotide repeat dynamics in <i>Dmt-D</i> cells arrested by apicidin.....	216
7.2.5. Trinucleotide repeat dynamics in confluent <i>Dmt-D</i> cell cultures exposed to interferon $\alpha$ .....	224
7.3. <i>Discussion</i> .....	224
<b>8. <i>Pms2</i> as a genetic modifier of trinucleotide repeat dynamics.....</b>	<b>230</b>
8.1. <i>Introduction</i> .....	230
8.2. <i>Results</i> .....	233
8.2.1. Breeding <i>Dmt-D</i> mice onto a <i>Pms2</i> -deficient genetic background.....	233
8.2.2. Mouse genotyping for the <i>Pms2</i> deletion.....	233
8.2.3. Somatic mosaicism in heterozygous <i>Pms2</i> knock-out mice.....	235
8.2.4. Somatic mosaicism in homozygous <i>Pms2</i> knock-out mice .....	239
8.2.5. Quantification of large repeat length mutations in a <i>Pms2</i> -deficient background. ....	241
8.3. <i>Discussion</i> .....	245
<b>9. Main conclusions, final discussion and future perspectives.....</b>	<b>250</b>
9.1. <i>Main conclusions</i> .....	250
9.1.1. Trinucleotide repeat dynamics and cell division .....	250
9.1.2. A possible model of trinucleotide repeat mutation.....	253
9.1.3. Dynamic equilibrium between MMR heteroduplexes and trinucleotide repeat mutation .....	256
9.1.4. Involvement of NER, BER and recombination events in triplet repeat mutation .....	257
9.1.5. Central involvement of DNA topology in the dynamics of expanded trinucleotide repeats ...	258
9.2. <i>On the molecular bases of tissue- and cell-specific somatic mosaicism</i> .....	259
9.3. <i>Future perspectives</i> .....	261
9.3.1. Novel therapeutical routes .....	261
<b>References.....</b>	<b>263</b>
<b>Publication.....</b>	<b>293</b>

# List of tables

<b>Table 1.1.</b>	Summary of trinucleotide repeat disorders .....	22
<b>Table 1.2.</b>	Protein localisation and patterns of degeneration in polyglutamine diseases.....	25
<b>Table 1.3.</b>	<i>mut</i> homologues in yeast and humans.....	46
<b>Table 1.4.</b>	Phenotypes of MMR-deficient mice .....	48
<b>Table 2.1</b>	Chemicals and reagents .....	62
<b>Table 2.2.</b>	Plastic materials and suppliers .....	63
<b>Table 2.3.</b>	Tissue culture materials and suppliers .....	63
<b>Table 2.4.</b>	Enzymes and suppliers .....	64
<b>Table 2.5.</b>	Nucleic acid size markers and/or mass ladders and their suppliers .....	64
<b>Table 2.6.</b>	Kits and suppliers.....	65
<b>Table 2.7.</b>	Oligonucleotides name, 5'-3' sequence, melting temperature ( <i>T<sub>m</sub></i> ) and target sequence.....	67
<b>Table 3.1.</b>	Annealing temperatures and number of cycles for each primer combination used in the <i>Dmt</i> mouse PCR genotyping analyses .....	92
<b>Table 3.2.</b>	Population doubling times of cell lines and relative trinucleotide repeat stability.....	93
<b>Table 4.1.</b>	Relative mobilities of CAG•CTG-containing PCR products through native polyacrylamide gels.....	122
<b>Table 6.1.</b>	Summary of the effects of multiple chemical treatments on the dynamics of expanded trinucleotide repeats in D2763Kc2 cells .....	190
<b>Table 6.2.</b>	Summary of the effects of multiple chemical treatments on the dynamics of expanded trinucleotide repeats in D4132K cells .....	192

# List of figures

<b>Figure 1.1.</b>	Location of trinucleotide repeat tracts in triplet repeat disease genes. ....	23
<b>Figure 1.2.</b>	Genomic organisation of the human <i>DMI</i> locus. ....	32
<b>Figure 1.3.</b>	Human DNA MMR protein complexes. ....	47
<b>Figure 1.4.</b>	DNA polymerase slippage model for trinucleotide repeat expansion. ....	51
<b>Figure 1.5.</b>	A model of repeat expansion via a FEN1-resistant folded flap. ....	54
<b>Figure 1.6.</b>	Expansion of trinucleotide repeats by double strand break repair. ....	57
<b>Figure 1.7.</b>	Mechanism for trinucleotide repeat expansion in germ cells by gap repair. ....	59
<b>Figure 2.1.</b>	Annealing sites for the oligonucleotides used to amplify the <i>Dmt162</i> transgene. ....	68
<b>Figure 3.1.</b>	Mouse genotyping for the <i>Dmt162</i> transgene. ....	92
<b>Figure 3.2.</b>	Growth dynamics of the D2763 cell cultures. ....	94
<b>Figure 3.3.</b>	Fibroblastic phenotype of D2763 cultured cells. ....	94
<b>Figure 3.4.</b>	Repeat length variation in <i>Dmt-D</i> cultured cells from a six-month-old mouse. ....	96
<b>Figure 3.5.</b>	Repeat length distributions in cultured kidney cells. ....	98
<b>Figure 3.6.</b>	Age-of-donor effects on repeat length variation in cultured kidney cells. ....	100
<b>Figure 3.7.</b>	Fibroblastic phenotype of cultured kidney cells from different aged mice. ....	102
<b>Figure 3.8.</b>	Repeat length variation in single cell clones of cultured kidney cells. ....	104
<b>Figure 3.9.</b>	Competition growth assays between single cell-derived clones. ....	106
<b>Figure 3.10.</b>	Trinucleotide repeat length variability in liver tissue samples from two 24-month-old <i>Dmt-D</i> mice. ....	108
<b>Figure 3.11.</b>	Fibroblastic phenotype of cultured kidney cells from a <i>Dmt-E</i> mouse. ....	110
<b>Figure 3.12.</b>	Trinucleotide repeat stability in cultured kidney cells from a <i>Dmt-E</i> mouse. ....	110
<b>Figure 4.1.</b>	Models of S-DNA and SI-DNA structures. ....	117
<b>Figure 4.2.</b>	Analysis of CAG•CTG-containing PCR products by native polyacrylamide gel electrophoresis. ....	120
<b>Figure 4.3.</b>	Increase mobility in native polyacrylamide gels as a function of repeat number. ....	121
<b>Figure 4.4.</b>	Percentage of B-DNA conformation as a function of repeat length. ....	121
<b>Figure 4.5.</b>	Reannealing experiments on PCR products containing CAG•CTG repeats. ....	124
<b>Figure 4.6.</b>	Generation of alternative DNA structures by PCR amplification. ....	125
<b>Figure 4.7.</b>	Assay for RecA activity on CAG•CTG repetitive sequences. ....	127
<b>Figure 4.8.</b>	Novobiocin treatment and expanded CAG•CTG repeat dynamics in D2763K cells. ....	130
<b>Figure 4.9.</b>	Novobiocin treatment and expanded CAG•CTG repeat dynamics in D2763Kc2 cells. ....	132
<b>Figure 4.10.</b>	Novobiocin treatment and expanded CAG•CTG repeat dynamics in D4132K cells. ....	134

<b>Figure 5.1.</b>	The bacterial 8-oxoG base excision repair system. ....	144
<b>Figure 5.2.</b>	Sensitivity of <i>Dmt-D</i> cultured cells to hydrogen peroxide.....	148
<b>Figure 5.3.</b>	Preliminary analysis of CAG•CTG repeat instability in <i>Dmt-D</i> kidney and eye cells treated with hydrogen peroxide.....	150
<b>Figure 5.4.</b>	Hydrogen peroxide treatment and expanded CAG•CTG repeat dynamics in D2763Kc2 cells.....	153
<b>Figure 5.5.</b>	Hydrogen peroxide treatment and expanded CAG•CTG repeat dynamics in D4132K kidney cells.....	155
<b>Figure 5.6.</b>	Hydrogen peroxide treatment and expanded CAG•CTG repeat dynamics in E3994K kidney cells.....	157
<b>Figure 5.7.</b>	Ethidium bromide treatment and expanded CAG•CTG repeat dynamics in D2763Kc2 cells. ...	159
<b>Figure 5.8.</b>	PCR amplification of mouse mitochondrial DNA extracted from <i>Dmt-D</i> cells treated with ethidium bromide. ....	160
<b>Figure 5.9.</b>	Quantification of <i>cytochrome oxidase II</i> mRNA levels in ethidium bromide treated cells. ....	162
<b>Figure 5.10.</b>	Effect of ethidium bromide on the mobility of expanded CAG•CTG sequences in agarose gels. 163	
<b>Figure 5.11.</b>	Effect of ethidium bromide incubation on the mobility of expanded CAG•CTG sequences through agarose gels. ....	164
<b>Figure 5.12.</b>	Effect of ethidium bromide on the mobility of CAG•CTG-containing plasmid DNA through agarose gels. ....	166
<b>Figure 5.13.</b>	Effect of ethidium bromide on the mobility of CAG•CTG-containing mouse genomic DNA fragments. ....	167
<b>Figure 5.14.</b>	Intracellular localisation of rhodamine-6G in mouse kidney cells. ....	169
<b>Figure 5.15.</b>	Rhodamine-6G treatment and expanded CAG•CTG repeat dynamics in D2763Kc2 cells.....	170
<b>Figure 5.16.</b>	Investigating the progressive effect of rhodamine-6G on expanded CAG•CTG repeat dynamics.....	172
<b>Figure 6.1.</b>	Sensitivity of <i>Dmt-D</i> cultured cells to UV-C light exposure. ....	184
<b>Figure 6.2.</b>	Experimental design to assess the effect of UV-C light exposure on trinucleotide repeat dynamics in <i>Dmt-D</i> cultured cells.....	184
<b>Figure 6.3.</b>	UV-C exposure and expanded CAG•CTG repeat dynamics in D2763K and D2763L cells. ....	186
<b>Figure 6.4.</b>	UV-C exposure and expanded CAG•CTG repeat dynamics in D2763Kc2 kidney cells.....	187
<b>Figure 6.5.</b>	Time course monitoring of the effects of UV-C radiation on the expanded CAG•CTG repeat dynamics in D2763Kc2 cells.....	189
<b>Figure 6.6.</b>	Caffeine treatment and expanded CAG•CTG repeat dynamics in D2763Kc2 kidney cells. ....	191
<b>Figure 6.7.</b>	Chemical treatment and expanded CAG•CTG repeat dynamics in D4132K cells. ....	193
<b>Figure 7.1.</b>	Trinucleotide repeat dynamics in <i>Dmt-D</i> kidney cells maintained at high density by serum starvation. ....	210
<b>Figure 7.2.</b>	Trinucleotide repeat dynamics in D4393K cells arrested by contact inhibition.....	212
<b>Figure 7.3.</b>	Immunocytochemical characterisation of D4393K kidney cells. ....	213
<b>Figure 7.4.</b>	BrdU incorporation analysis on D4393K cells arrested by contact inhibition. ....	213
<b>Figure 7.5.</b>	Dynamics of expanded CAG•CTG trinucleotide repeats in D2763Kc2 kidney cells arrested by mitomycin C.....	215
<b>Figure 7.6.</b>	BrdU incorporation analysis on D2763Kc2 kidney cells arrested by mitomycin C or apicidin exposure. ....	217

<b>Figure 7.7.</b>	Repeat distributions in D2967K cells arrested by mitomycin C.....	218
<b>Figure 7.8.</b>	Repeat distributions in D3111K cells arrested by mitomycin C.....	219
<b>Figure 7.9.</b>	Morphological changes in <i>Dmt-D</i> kidney cells cultured in 320 nM apicidin.....	221
<b>Figure 7.10.</b>	BrdU incorporation analysis on D979K cells arrested by apicidin.....	221
<b>Figure 7.11.</b>	Dynamics of expanded CAG•CTG trinucleotide repeats in D2763Kc2 arrested by apicidin. ...	222
<b>Figure 7.12.</b>	Dynamics of expanded CAG•CTG trinucleotide repeats in D979K cells arrested by apicidin. ....	223
<b>Figure 7.13.</b>	Dynamics of expanded CAG•CTG trinucleotide repeats in D979K cells arrested by interferon $\alpha$ .....	225
<b>Figure 8.1.</b>	Mouse <i>Pms2</i> gene disruption by insertion of a <i>neo</i> resistance cassette.....	234
<b>Figure 8.2.</b>	Mouse <i>Pms2</i> genotyping by PCR analysis. ....	234
<b>Figure 8.3.</b>	Somatic mosaicism in 24-month-old <i>Dmt-D</i> mice heterozygous for the <i>Pms2</i> deficiency. ....	236
<b>Figure 8.4.</b>	Quantification of the effect of the disruption of one <i>Pms2</i> allele on the levels of trinucleotide somatic mosaicism detected in the lung and kidney of <i>Dmt-D</i> mice. ....	237
<b>Figure 8.5.</b>	Somatic mosaicism in 13-month-old <i>Dmt-D</i> mice heterozygous for the <i>Pms2</i> deletion. ....	238
<b>Figure 8.6.</b>	Somatic mosaicism in lung and kidney of an eight-month-old <i>Dmt-D</i> mouse homozygous for the <i>Pms2</i> deletion. ....	239
<b>Figure 8.7.</b>	Quantification of the effect of a homozygous <i>Pms2</i> -deficient genetic background on the trinucleotide somatic mosaicism detected in the lung and kidney of <i>Dmt-D</i> mice. ....	240
<b>Figure 8.8.</b>	Somatic mosaicism in six-month-old <i>Dmt-D</i> mice homozygous for the <i>Pms2</i> deletion. ....	242
<b>Figure 8.9.</b>	High DNA input SP-PCR amplifications for the quantification of large repeat number changes in the heart and brain of a <i>Pms2</i> -null homozygous mouse. ....	243
<b>Figure 8.10.</b>	Frequency of large trinucleotide repeat number mutants in the brain, heart, lung and kidney in a <i>Pms2</i> -deficient background.....	244
<b>Figure 9.1.</b>	Model of somatic trinucleotide repeat expansion mediated by MMR proteins.....	254

# Acknowledgements

I promise I won't thank everybody I've ever met in my entire life, but right now, when this journey seems to be finally coming to an end, it feels like I really should. There're so many people who, in one way or another, have helped, encouraged, supported or even insulted me over the years, and they all deserve a mention here. I know I'm taking the serious risk of sounding soppy, but quite frankly I couldn't care less. These few pages will certainly be the only bit of this thesis that many of you will ever read (I know what I'm talking about, I've done the same). And most of those who'll take up the adventure of keeping on reading through the rest, will only probably remember, comment and be cheeky about these few pages (I've done it too). So let's give you all a great opportunity to be cheeky.

It's been quite a story so far (I know it sounds pretentious; there's no need to tell me...). A journey that started some time ago when I joined a lab in Porto, for my undergraduate honours project. I was lucky enough to end up working with the sweetest supervisor somebody can wish for, Rosário: my first boss ever. I'd like to thank you for your patience, kindness, for all the valuable things you taught me, and above all for your encouragement and advice. I then moved on, and joined the Gulbenkian PhD Programme in Biology and Medicine, now in Lisbon. I should thank António Coutinho and Alexandre Quintanilha for having the faith to give me the wonderful chance to join the Gulbenkian PhD Programme. And also Paulo Vieira, our "manager" at the Gulbenkian Institute for Science, Manuela Cordeiro and Helena Matias for having to deal with all the boring beaurocratic stuff. Thanks ever so much!

One year later, after a short visit to Glasgow, I was moving to Scotland to start working on my PhD project. This would have never happened without the vision, the ideas and the enthusiasm of my supervisor, Darren. I'll always remember the day when we both mashed up some tissue samples together, without really knowing what we were supposed to be doing, and that was when we started growing the very first cells ever... It turned out to be the beginning of an exciting and enjoyable story. Thanks a lot for everything, Darren, especially for your endless, endless support, patience and encouragement throughout my PhD.

I'm grateful to my assessors, Mark Bailey and Marshal Stark, for the very helpful discussions and the long "chats" we had together during the course of my project.

To Maria João Saraiva, my Portuguese supervisor. Thanks for being so supportive and involved over the years, despite the famous "e-mail faux pas". Thanks for all your tips, pieces of advice and wise comments, and for your visit to Glasgow (I'm sorry about the weather, by the way... but I'm afraid that's Glasgow, and it doesn't get any better).

When I first started working on my PhD project it felt like I was facing an endless task ahead of me. But now it feels like it was all done and over within a couple of months (not exactly, but anyway...), thanks to all the nice people I've been working with in Anderson College, Level 5,

and to the great friends I've made here, who made me feel welcome and made me laugh over the last years. I can honestly say I started working with a bunch of strangers, and ended up working with a bunch of friends. I'd like to thank to the DGM team, especially John McAbney for taking care of us, for making up many, many solutions and buffers, dealing with the reps and, strangely enough, never complaining about it. Instead, you can still find the energy to keep us entertained with your CD player and your cracking jokes. Also Christine and Jon, for making that corner of the lab so special... I must admit there were times when I was helped out by three talented and keen undergraduate students: Gareth Barnes, Laura Ingram and Sanam Mustafa. Thanks for giving me a hand guys! I hope I wasn't too grumpy and moody at times.

I'd particularly like to express my most sincere gratitude to those who, in one way or another, looked after me when I first got to this strange city: Catherine and Graham (you were superb), Sue (the best landlady ever) and Darren. A very special "thanks" goes to my partner in crime Laura, and also to Giorgia, Alan, Christine, Shirley, Saadia, Ann and Mei for all the laughter and pints of lager you bought me. And to Jeff, for putting up with my bad mood over the last few months.

But above all, I'd like to thank those I left behind in Portugal, but who made me feel like we were just a stone throw away from each other, with their e-mails, phone calls, letters and cards. You know who you are, but I'm afraid I'll have to mention some names here. May the others forgive me. Liliana, Mónica, Cláudia and André. You're the best, and you deserve everything you've ever wished for, twice and over again. Thanks for always being in touch, for listening and for all the time and stories we've shared together. It's been great fun so far! Just like the song goes (by someone who I'm too embarrassed to write it here): "You've helped me holding my head up high through the storm, walking on through the wind and the rain; I've never walked alone".

To my family, that big bunch of mothers and fathers, uncles and aunties, grandpas and grannies, I've got back home. To my mum Teresa and my dad Manel: thanks for bringing me up the way you did, teaching me to always follow my dreams, and never be afraid of taking the chance. Thanks for the support you've given me all the way, for trusting and believing in me. And finally, and most importantly, thanks for loving me enough to let me go away. I'll be back one day, promise. To my Avô Jaime, my lucky star, for your constant support, enthusiasm and friendship. And Titi and Fonfom, for some of the best times of my life, which, after all, I spent with you, in those long and magic summer holidays. This is for you all!

Finally, I'm also in debt to Fundação Calouste Gulbenkian (Portugal) and Fundação para a Ciência e Tecnologia (Portugal) for paying me to do my job over the last few years.

**To my family and friends**

*Esta palavra saudade  
Sete letras de ternura  
Sete letras de ansiedade  
E outras tantas de aventura*

*Esta palavra saudade  
A mais bela e a mais pura  
Sete letras de verdade  
E outras tantas de loucura*

*Esta palavra saudade  
Sabe a sumo de limão  
Tem o travo da amargura  
Que nasceu do coração*

*Esta palavra saudade  
Sabe ao gosto das amoras  
Cada vez que tu não vens  
Cada vez que tu demoras*

*Palavra amarga e doce  
Debruçada na idade  
Palavra como se fosse  
Um resto de mocidade*

*Marcada por sete letras  
A ferro e fogo no tempo  
Palavra dos poetas  
Que disparam contra o vento*

José Carlos Ary dos Santos

The research reported in this thesis is my own original work, except where otherwise stated, and has not been submitted for any other degree.

Mário Gomes-Pereira

June 2002

# List of Abbreviations

°C	Degrees Celsius
[ $\alpha$ - <sup>32</sup> P]dCTP	$\alpha$ - <sup>32</sup> P-labelled 2'-deoxycytidine-5'-triphosphate
$\lambda$	Wavelength
$\mu$	Micro ( $10^{-6}$ )
6-4PP	Pyrimidine-pyrimidone 6-4 photoproduct
8-oxoG	7,8-Dihydro-2'-deoxiguanosine
AP	Abasic
APS	Ammonium persulphate
AraC	1- $\beta$ -D-arabinofuranosyl-cytosine
ATM	Ataxia-telangiectasia mutated
ATR	ATM and Rad3-related
b	Base
BER	Base excision repair
bp	Base pair
BrdU	5'-Bromo-3-deoxyuridine
BSA	Bovine serum albumin
CAG	Trinucleotide of cytosine, adenosine and guanine
CCTG	Tetranucleotide of cytosine, cytosine, thymine and guanine
CDK	Cyclin dependent kinase
cDNA	Complementary deoxyribonucleic acid
CGG	Trinucleotide of cytosine, guanine and guanine
Ci	Curie
CPD	Cyclobutane-pyrimidine dimer
cpm	Counts per minute
cps	Counts per second
CTG	Trinucleotide of cytosine, thymine and guanine
CUG	Trinucleotide of cytosine, uracil and guanine
CUG-BP	CUG-binding protein
dATP	2'-Deoxyadenosine-5'-triphosphate
dCTP	2'-Deoxycytidine-5'-triphosphate
DEPC	Diethylpyrocarbonate
dGTP	2'-deoxyguanosine-5'-triphosphate
DM	Myotonic dystrophy, <i>dystrophia myotonica</i>
DMEM	Dulbecco's modified Eagle medium
DMPK	Dystrophia myotonica protein kinase
DMSO	Dymethylsulphoxide
DMWD	Dystrophia myotonica-containing WD repeat motif

DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotidetriphosphate
DRPLA	Dentatorubral pallidoluyasian atrophy
DTT	Dithiotheitol
dTTP	2'-Deoxythymidine-5'-triphosphate
<i>E. coli</i>	<i>Escherichia coli</i>
EBV	Epstein-Barr virus
EDTA	Ethylenediaminetetracetic acid
EtBr	Ethidium bromide
EXP	Triplet repeat expansion proteins
FBS	Foetal bovine serum
FEN1	Flap structure-specific endonuclease 1
FITC	Fluoresceine isothiocyanate isomer I
FMR	Fragile X mental retardation
FMRP	Fragile X mental retardation protein
FRAXA	Fragile X syndrome
FRAXE	Fragile X site E
FRDA	Friedreich ataxia
<i>g</i>	Gravity acceleration
<i>g</i>	Gram
GAA	Trinucleotide of guanine, adenine and adenine
HD	Huntington disease
HNPCC	Hereditary non-polyposis colorectal cancer
IPTG	Isopropylthyo-β-D-galactoside
<i>k</i>	Kilo (10 <sup>3</sup> )
KLHL1	Kelch-like 1
<i>l</i>	Litre
LB	Luria Bertani
LBCL	Lymphoblastoid cell lines
<i>m</i>	Milli (10 <sup>-3</sup> )
<i>M</i>	Molar
MDR	Multidrug resistance
MED1	Methyl-CpG binding endonuclease 1
MLH	MutL homologue
MMR	Mismatch repair
MOPS	3-( <i>N</i> -Morpholino)-propanesulphonic
mRNA	Messenger ribonucleic acid
MSH	MutS homologue
mUSF	Mouse upstream stimulatory factor
MW	Molecular weight
<i>N</i>	Nucleotide of adenine, cytosine, guanine or thymidine.

n	Nano ( $10^{-9}$ )
<i>neo</i>	Neomycin
NER	Nucleotide excision repair
NMR	Nuclear magnetic resonance
NSAID	Nonsteroidal anti-inflammatory drug
OD	Optical density
OGG	7,8-Dihydro-2'-deoxyguanosine DNA glycosylase
p	Pico ( $10^{-12}$ )
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCNA	Proliferating-cell nuclear antigen
PCR	Polymerase chain reaction
PIKK	Phosphoinositide 3-kinase related kinase
PKR	Double-stranded RNA-activated protein kinase
PMS	Post-meiotic segregation
PMSF	Phenylmethylsulphonyl fluoride
PPP2R2B	Protein phosphatase 2A, regulatory subunit B
RNA	Ribonucleic acid
ROS	Reactive oxygen species
rRNA	Ribosomal ribonucleic acid
RSHL	Radial spokehead-like
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SBMA	Spinal and bulbar muscular atrophy
SCA	Spinocerebellar ataxia
S-DNA	Slipped-stranded deoxyribonucleic acid
SDS	Sodium dodecyl sulphate
SI-DNA	Slipped-stranded intermediate deoxyribonucleic acid
SIX5	Sine oculis related homeobox 5
SP-PCR	Small pool polymerase chain reaction
SV40	Simian virus 40
TCR	Transcription-coupled repair
TEMED	<i>NNN'N'</i> -Tetramethylethylenediamine
Tris	Tris(hydroxymethyl)amino methane
TRITC	Tetramethylrhodamine isothiocyanate
U	Unit
UTR	Untranslated region
UV	Ultraviolet
v	Volume
V	Volt
w	Weight
X-gal	5-Bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside

# 1. Introduction

## 1.1. Trinucleotide repeats and human disease

In 1991, a new mutational disease mechanism was identified as the molecular basis of two human genetic diseases, fragile X syndrome (FRAXA), and spinobulbar muscular atrophy (SBMA): the expansion of simple trinucleotide repeats. A DNA repetitive trinucleotide sequence (GGC•CCG or CAG•CTG), was found to be increased in copy number beyond a certain threshold in affected individuals (Fu *et al.*, 1991; La Spada *et al.*, 1991). Triplet repeat expansions have since been found to cause an ever-increasing list of human diseases (Table 1.1), all showing neurological and/or neuromuscular involvement, and clinical pictures that range from mild to severely debilitating, or even fatal. Trinucleotide repetitive sequences are usually polymorphic in normal individuals, being usually shorter than 35 repeats, depending on the loci (Cummings and Zoghbi, 2000b). However, once into the expanded disease-associated range (usually corresponding to more than 35 triplet repeat units) the repeat tracts become dramatically unstable, exhibiting very high mutation rates, with a marked tendency for further expansion, not only in the germline but also in somatic cells (Richards, 2001).

Following the description of FRAXA and SBMA mutations, came the identification of five other neurological diseases caused by untranslated triplet repeats: myotonic dystrophy type 1 (DM1) (Aslanidis *et al.*, 1992; Brook *et al.*, 1992; Buxton *et al.*, 1992; Fu *et al.*, 1992; Harley *et al.*, 1992; Mahadevan *et al.*, 1992), fragile XE mental retardation (FRAXE) (Knight *et al.*, 1993), Friedreich ataxia (FRDA) (Campuzano *et al.*, 1996), spinocerebellar ataxia (SCA) type 8 (SCA8) (Koob *et al.*, 1999) and SCA12 (Holmes *et al.*, 1999). Additionally, other neurodegenerative diseases were described as the result of an expansion of CAG•CTG repeats within multiple genes, encoding for polyglutamine stretches in the corresponding proteins: Huntington disease (HD) (The Huntington's Disease Collaborative Research Group, 1993); dentatorubral pallidoluysian atrophy (DRPLA) (Koide *et al.*, 1994; Nagafuchi *et al.*, 1994), SCA1 (Orr *et al.*, 1993), SCA2 (Sanpei *et al.*, 1996), Machado-Joseph disease (MJD), also termed SCA3 (Cancel *et al.*, 1995), SCA6 (Zhuchenko *et al.*, 1997), SCA7 (David *et al.*, 1997) and SCA17 (Koide *et al.*, 1999; Nakamura *et al.*, 2001). Trinucleotide repeat disease can therefore be grouped into two subclasses based on the relative location of the trinucleotide repeat within the affected gene (Figure 1.1). The first subclass, collectively referred to as “polyglutamine diseases”, is characterised by exonic CAG•CTG repeats that code for polyglutamine tracts. The second subgroup includes disorders caused by repeats that map within non-coding regions of a gene. At the moment all these human conditions have very limited treatment options.

Both FRAXA and FRAXE are associated with fragile sites, which are chromosomal points that show a high frequency of non-random breaks, under specific *in vitro* conditions. Fragile sites are readily observed by cytogenetic analysis, following exposure of the cells, from which the chromosomes are prepared, to particular conditions of tissue culture or chemical agents (Sutherland *et al.*, 1998; Sutherland and Richards, 1995).

**Table 1.1. Summary of trinucleotide repeat disorders** (adapted from Sinden *et al.*, 2002)

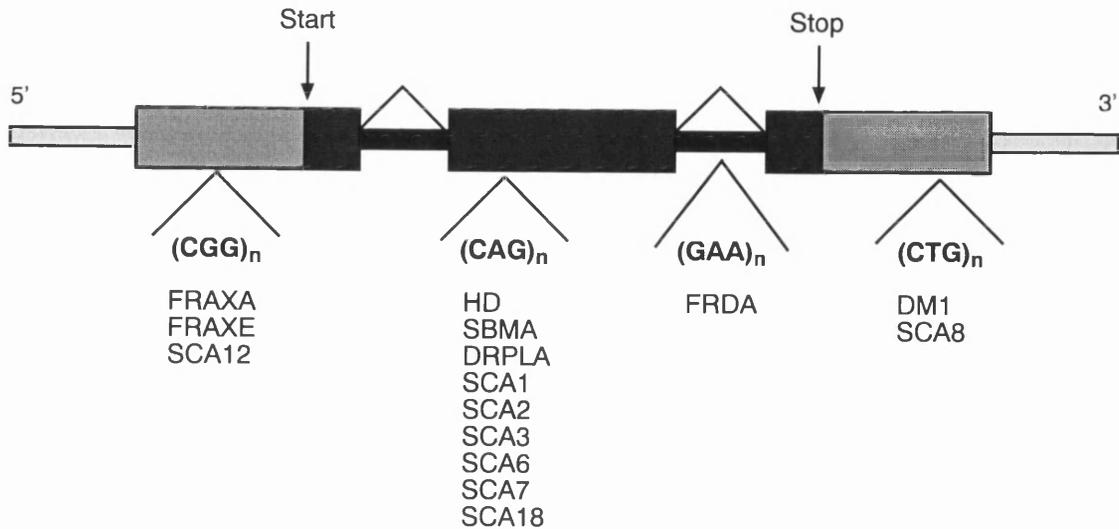
Disease	Inheritance	Locus	Gene	Protein product	Repeat sequence <sup>a</sup>	Repeat location
<b>DM1</b>	Autosomal dominant	19q13.3	<i>DMPK</i>	Myotonic dystrophy protein kinase	CTG	3'UTR
<b>DRPLA</b>	Autosomal dominant	12p13.31	<i>DRPLA</i>	Atrophin-1	CAG	Coding
<b>FRAXA</b>	X-linked recessive	Xq27.3	<i>FMR1</i>	Fragile X mental retardation protein	CGG	5'UTR
<b>FRAXE</b>	X-linked recessive	Xq28	<i>FMR2</i>	FMR2 protein	CGG	5'UTR
<b>FRDA</b>	Autosomal recessive	9q13-21.1	<i>X25</i>	Frataxin	GAA	Intronic
<b>HD</b>	Autosomal dominant	4p16.3	<i>HD/IT15</i>	Huntingtin	CAG	Coding
<b>SBMA</b>	X-linked sex-limited	Xq13-21	<i>AR</i>	Androgen receptor	CAG	Coding
<b>SCA1</b>	Autosomal dominant	6p23	<i>SCA1</i>	Ataxin-1	CAG	Coding
<b>SCA2</b>	Autosomal dominant	12q23.1	<i>SCA2</i>	Ataxin-2	CAG	Coding
<b>SCA3/MJD</b>	Autosomal dominant	14q32.1	<i>SCA3/MJD</i>	Ataxin-3	CAG	Coding
<b>SCA6</b>	Autosomal dominant	19p13	<i>CACNA1A</i>	$\alpha_{1A}$ -voltage-dependent calcium channel subunit	CAG	Coding
<b>SCA7</b>	Autosomal dominant	3p12-13	<i>SCA7</i>	Ataxin-7	CAG	Coding
<b>SCA8</b>	Autosomal dominant	13q21	None	None	CAG	3' end <sup>b</sup>
<b>SCA12</b>	Autosomal dominant	5q31-33	<i>PPP2R2B</i>	Protein phosphatase 2, regulatory subunit $\beta$	CAG	5'UTR
<b>SCA17</b>	Autosomal dominant	6q27	<i>TBP</i>	TATA-binding protein	CAG	Exonic

<sup>a</sup> On the coding/sense DNA strand

<sup>b</sup> Untranslated transcript (see Section 1.2.2.5)

Not all trinucleotide expansions (>35 repeats) are associated with a clinical phenotype, since at least two non-pathogenic CAG•CTG repeat expansions have been described (Breschel *et al.*, 1997; Ikeuchi *et al.*, 1998; Nakamoto *et al.*, 1997).

As a consequence of the expanding number of neurological and neuromuscular diseases associated with trinucleotide repeat sequences, both the molecular mutational mechanisms and the pathogenic pathways leading from mutation to disease phenotype, have been the subjects of intense research. The common properties of the repeats in different diseases and fragile sites have given insight into this unique form of DNA instability. Referring to the idiosyncratic properties of trinucleotide repeat sequences, the term “dynamic mutation” was introduced to distinguish mutant alleles, which once beyond a specific size threshold become dramatically unstable in a length-dependent manner (Richards and Sutherland, 1992).



**Figure 1.1. Location of trinucleotide repeat tracts in triplet repeat disease genes.**

An idealised gene is shown in the figure. The dark grey areas represent the coding region that is transcribed and translated into the final gene product. The transcription start and stop positions are indicated. The light grey bars represent the 5' upstream and 3' downstream untranslated regions (UTR). Introns are shown in black. Both disease-associated  $(CGG)_n$  triplet repeats are located in the 5' UTR. The  $(CAG)_n$  repeat in SCA12 is also located in the 5' UTR. Glutamine encoding  $(CAG)_n$  expansions are located within exons of the gene coding regions. The  $(GAA)_n$  repeat in FRDA is located in an intron. The non-coding  $(CTG)_n$  repeats in DM1 and SCA8 map within the 3' UTR of the associated genes (adapted from Sinden *et al.*, 2002).

The application of comparative molecular anatomy to the study of molecular fragile sites and disease loci associated with triplet repeat disorders has led to the identification of a number of common properties, which most likely reflect common molecular mechanisms for both the genesis of expanded alleles and pathophysiology. First, the mutant repeats exhibit both somatic and germline instability, with a clear bias towards expansion in repeat copy number, and with mutation rates related to the initial number of repeats. While polyglutamine disorders typically involve smaller repeat expansions (ranging from ~35 to ~100 units) the second subclass of conditions, caused by non-coding repeats, is usually characterised by massive repeat expansions (from ~35 up to several hundreds, or even thousands of repeats), both between generations and in the soma. Second, rare founder events hint at the existence of alleles with increased likelihood of undergoing changes in repeat copy number. Third, diseases caused by trinucleotide repeat expansion exhibit a relationship between copy number of the repeat and the severity and age of onset of symptoms (Richards, 2001). Taken together, these properties account for the phenomenon of clinical anticipation, the increasing severity, more rapid disease progression and decreasing age of onset in successive generations within an affected family. Although there are some unifying phenomena connecting the dynamic mutation disorders, little is understood, not only about the pathways that lead from mutation to disease, but also about the molecular mechanisms of repeat mutation (Cummings and Zoghbi, 2000a; Richards, 2001; Sinden, 2001).

## 1.2. Molecular pathogenesis

Increasing the copy number of an existing DNA repeat sequence might appear to be a relatively simple process with no major biological implications. In reality, the biological consequences can be remarkably complex. Indeed, trinucleotide repeat expansions may cause pathology through a variety of mechanisms, including interference with DNA structure, transcription, RNA-protein interaction and altered protein conformation and metabolism. This is clearly indicated by the fact that the currently known triplet repeat expansions associated with disease are found in both 5' and 3' untranslated regions (UTR), introns and within coding sequences of various affected genes (Cummings and Zoghbi, 2000b). A brief summary of the molecular pathophysiological events that may link trinucleotide repeat expansion to the disease phenotype is presented below.

### 1.2.1. Polyglutamine diseases

Polyglutamine diseases are most frequently autosomal dominant, with the exception of spinobulbar muscular atrophy (SBMA), which exhibits an X-linked sex-limited pattern of inheritance (Table 1.1). Although the proteins mutated in these conditions do not share any degree of identity or similarity, aside from the polyglutamine tract, several outstanding features are shared by this subclass of disorders, hinting at a possible shared mechanism of pathogenesis. All these diseases are progressive, typically striking in mid-life and causing increasing dysfunction and eventual neuronal loss 10 to 20 years after the onset of symptoms (Cummings and Zoghbi, 2000a; Cummings and Zoghbi, 2000b; Zoghbi and Orr, 2000). Analysis of repeat size and symptomatology in polyglutamine disease patients has demonstrated an inverse relationship between the repeat size and the age at onset. Each disorder shows a characteristic threshold for glutamine tract length, below which symptoms do not occur. Above the threshold, the progressive decrease in age of onset with increasing polyglutamine length shows a slightly different slope for each disorder, indicating that the increased severity, due to each extra glutamine residues, depends on the protein context. It is also clear from careful study of various kindreds that additional factors, probably either environmental or genetic, contribute to the onset of disease (Gusella and MacDonald, 2000; Zoghbi and Orr, 2000). Despite the widespread expression of the relevant protein throughout the brain and other tissues, only a certain subset of neurons is vulnerable to dysfunction in each of these diseases (Table 1.2). The variability in cell-specific degeneration is lost, however, when the expansions are very large, leading to severe juvenile-onset disease. In these cases, there is significant overlap in the phenotypes (Cummings and Zoghbi, 2000b). The loss of cell specificity hints that toxicity is probably much more widespread through neuronal and non-neuronal cells, which are normally spared when the repeat sizes are moderately expanded. It may also reinforce the idea that repeat expansions are smaller in the coding repeat disorders because of selective pressure against very large expansions, which are likely to be embryonic lethal. Very long polyglutamine expansion may be so neurotoxic that they override their specific disease context,

causing a more widespread “glutaminopathy” (Cummings and Zoghbi, 2000a; Cummings and Zoghbi, 2000b; Paulson, 1999; Zoghbi and Orr, 2000).

**Table 1.2. Protein localisation and patterns of degeneration in polyglutamine diseases** (adapted from Zoghbi and Orr, 2000).

Disease	Protein	Normal protein localisation	Brain regions most affected
DRPLA	Atrophin-1	Cytoplasmic	Cerebellum, cerebral cortex, basal ganglia, Luys body
HD	Huntingtin	Cytoplasmic	Striatum, cerebral cortex
SBMA	Androgen receptor	Nuclear and cytoplasmic	Anterior horn and bulbar neurons, dorsal root ganglia
SCA1	Ataxin-1	Nuclear in neurons	Cerebellar Purkinje cells, dentate nucleus, brain stem
SCA2	Ataxin-2	Cytoplasmic	Cerebellar Purkinje cells, brain stem, fronto-temporal lobes
SCA3/MJD	Ataxin-3	Cytoplasmic	Cerebellar Purkinje cells, brain stem, spinal cord
SCA6	$\alpha_{1A}$ -voltage-dependent calcium channel subunit	Cell membrane	Cerebellar Purkinje cells, dentate nucleus, inferior olive
SCA7	Ataxin-7	Nuclear	Cerebellum, brain stem, macula, visual cortex
SCA17	TATA-binding protein	Nuclear	Cerebellar Purkinje cells, molecular layer, dentate nucleus

### 1.2.1.1. Toxicity of polyglutamine tracts

There are no described cases of polyglutamine diseases caused by deletions or point mutations, and the CAG•CTG repeat expansion does not appear to compromise the normal function of the protein (Dragatsis *et al.*, 2000; Zoghbi and Orr, 2000), suggesting that these disorders do not result from a loss of gene function. A model of pathogenesis posits that the expanded glutamine tract mediates some undefined toxic gain-of-function, which results in neuronal dysfunction and death. Studies of mouse models have been the most compelling to show the enhanced toxicity of longer polyglutamine tracts in causing disease (Burrigh *et al.*, 1995; Ikeda *et al.*, 1996). The development of neuropathological phenotype by transgenic mice expressing a highly expanded polyglutamine tract (239 glutamines) alone, under the control of the human *androgen receptor* promoter (Adachi *et al.*, 2001), strongly suggested that polyglutamine tracts are toxic *per se*, and can induce neuronal dysfunction, irrespective of the protein context. In contrast, expression of an untranslated human *HD* transcript in transgenic mice, failed to cause the development of any phenotype despite the high levels of expression (Goldberg *et al.*, 1996). The causative agent in these conditions is clearly different from other diseases, in which untranslated triplet expansions are located in non-coding regions of mRNA transcripts. These findings

established that these inherited neurodegenerative disorders may appropriately be termed “polyglutamine diseases”.

### 1.2.1.2. Polyglutamine protein intranuclear inclusions

The presence of intranuclear inclusions is a shared pathological hallmark of polyglutamine disorders. These aggregates were first identified in mouse models of HD and in human HD brains (Becher *et al.*, 1998; Davies *et al.*, 1997; DiFiglia *et al.*, 1997), but they have also been described in other polyglutamine diseases. The predominant nuclear localisation of the inclusions is intriguing given the various subcellular localisations of the soluble forms of these proteins (Table 1.2). Nuclear inclusions might compromise neuronal function by sequestering transcription factors, or other regulatory nuclear proteins, containing glutamine-rich or pure polyglutamine domains, thereby altering transcription of genes that are critical for neuronal functions, mRNA splicing, or the export of proteins and RNA to the cytoplasm (Green, 1993; Paulson, 1999; Perutz *et al.*, 1994). The redistribution of proteasome components molecular chaperones to intranuclear inclusions suggests that a stress response is mounted to cope with the presence of toxic misfolded polyglutamine proteins (Davies *et al.*, 1999; Matilla *et al.*, 2001; Paulson, 1999).

The discovery of intranuclear inclusions suggested a common pathogenic pathway for all polyglutamine disorders. However, the initial evidence that nuclear inclusions were only present in vulnerable neurons, which once offered a clue to the selective vulnerability observed in polyglutamine diseases, is not holding true. Studies of brains of HD and SCA7 patients revealed that the regional distribution of the nuclear inclusions was not selectively restricted to specific regions of pathology (Gutkunst *et al.*, 1999; Holmberg *et al.*, 1998). Furthermore, studies in cell culture and mouse models of polyglutamine diseases, have shown that, while polyglutamine localisation in the nucleus is critical for pathogenesis (Klement *et al.*, 1998), the formation of nuclear inclusions does not correlate with neuronal death (Cummings *et al.*, 1999; Klement *et al.*, 1998; Saudou *et al.*, 1998). Consequently, it has been proposed that, rather than being pathogenic, nuclear inclusions may instead be protective against the toxic effects of the expanded polyglutamine (Sisodia, 1998). Soluble mutant proteins may actually be more toxic if not properly ubiquitinated, turned-over or possibly sequestered to a nuclear inclusion.

### 1.2.1.3. Cytoplasmic aggregates

It has become clear that, apart from intranuclear inclusions, cytoplasmic protein aggregates are also present in the cortex and striatum of adult and juvenile HD patients, but not in control brains (DiFiglia *et al.*, 1997). Cytoplasmic aggregates are typically more common than those in the nucleus, appear to accumulate with the duration of the disease and, most significantly, were observed in a presymptomatic HD patient (Gutkunst *et al.*, 1999). The development of cytoplasmic aggregates correlates with the progression of symptoms in transgenic mice (Li *et al.*, 1999; Schilling *et al.*, 1999), being preferentially detected in the striatal neurons, which degenerate during the early stages of HD knock-in mice, expressing full-length mutant huntingtin (Li *et al.*,

2001). Cytoplasmic polyglutamine aggregates may play a role in polyglutamine-induced neurodysfunction, as they promote a pathological mechanism of protein-protein interaction, which may result in recruitment of caspases, which upon activation may enter the nucleus and trigger the apoptotic cell death pathway (Li *et al.*, 2000c; Ona *et al.*, 1999; Sanchez *et al.*, 1999). Despite the latest progress, it still remains unclear whether protein aggregation does indeed contribute to pathogenesis, or whether it is simply a by-product of the disease process, indicative of cell stress.

#### 1.2.1.4. Cell dysfunction versus cell death

Although it had long been assumed that the neurological phenotype in polyglutamine patients resulted from neuronal death, the generation of transgenic mice expressing truncated forms of polyglutamine-containing peptides, revealed minimal and delayed neuronal death, relative to the development behavioural symptoms (Davies *et al.*, 1997; Mangiarini *et al.*, 1996). Gene targeting approaches enabled the generation of knock-in HD mouse models, carrying a precise insertion of a CAG•CTG expanded repeat into the appropriate position of the endogenous *Hdh* mouse gene, homologous to the human *HD* gene (Shelbourne *et al.*, 1999; Wheeler *et al.*, 1999b; White *et al.*, 1997). Some of these mice show aggressive behaviour (Shelbourne *et al.*, 1999) and impaired synaptic plasticity in the absence of neurodegeneration (Usdin *et al.*, 1999). Others exhibit disruption of striatal cell homeostasis and activation of cellular stress pathways (Trettel *et al.*, 2000), in the absence of obvious neuronal loss or gliosis (Menalled and Chesselet, 2002). Neuronal dysfunction, rather than neuronal death, may therefore be responsible for the pathological phenotype. Transcriptional alterations, prior to behavioural changes and cell death, could affect signalling pathways and neurotransmitter receptor levels (Cha *et al.*, 1998; Iannicola *et al.*, 2000; Lin *et al.*, 2000a; Luthi-Carter *et al.*, 2000), thereby providing the trigger for cell dysfunction and pathology in the absence of neuronal loss

#### 1.2.1.5. Cell-specific vulnerability in polyglutamine diseases

Cell specificity in polyglutamine diseases is an interesting, yet puzzling, feature of these conditions. If polyglutamine tracts are a necessary and sufficient condition to cause pathology (Adachi *et al.*, 2001; Marsh *et al.*, 2000; Ordway *et al.*, 1997), which factors determine the regional patterns of neurodegeneration, given that at least some of these proteins are widely expressed? The cell specificity of the various phenotypes could be due to a number of factors, such as cell-specific protein interactions and cell specificity of putative modifying proteins. Another possibility is that selective processing of full-length mutant protein is specific for a particular population of neurons, in agreement with the observation that N-terminal huntingtin fragments were preferentially found in striatal neurons of HD knock-in mice (Li *et al.*, 2000a). The distribution of neurologic symptoms may also depend on expression levels, as suggested by widespread development of neurological disease in mice expressing a truncated, but highly expanded, androgen receptor under the control of the prion promoter, in contrast with the motor symptoms developed when the expression of the same transgene was controlled by the neurofilament light chain promoter (Abel *et al.*, 2001). The

cell type specificity suggests that there may be intrinsic mechanisms in each cell, which render some cells resistant, while others become sensitive.

Interestingly, assessment of trinucleotide repeat dynamics in a knock-in HD mouse model revealed cell-specific expansion of the unstable CAG•CTG repeat in the striatum, the major affected tissue in HD (Kennedy and Shelbourne, 2000). This result suggests that ongoing somatic mutation may contribute to disease, by the continuous accumulation of expanded repeats. The time taken for the disease to manifest could therefore represent how long it takes for the disease-associated allele to reach a critical higher copy number. It is possible that over time, the striatal cells, no longer able to cope with fast accumulating mutant polyglutamine proteins, eventually succumb.

### **1.2.2. Non-coding repeat disorders**

Diseases in this subgroup are characterised by a complexity of symptoms and phenotypic variability. The mechanism of pathogenesis in non-coding repeat conditions varies from disease to disease, depending on the consequences of loss-of-function of the respective protein or, in some cases, acquired function of a toxic triplet repeat transcript. The trinucleotide sequence and its location within the affected gene may play a prominent role in dictating the unique mechanism of pathogenesis for each disease. Despite these inherent differences, many similarities exist within this group of diseases. First, the size and variation of the repeat expansion are much greater in the non-coding repeat disorders than in polyglutamine repeat conditions. Second, the non-coding repeat diseases are typically multisystemic disorders, involving the dysfunction and/or degeneration of many different tissues. Phenotypes within a disorder of this subclass are often variable, perhaps due to a more prominent degree of somatic repeat size heterogeneity in the non-coding repeats compared with the exonic repeat tracts; this is especially evident in FRAXA and DM1. Finally, many of the non-coding repeat disorders can be associated with a small pool of clinically silent, intermediate-sized expansions or premutations that may expand to the full mutation after germline transmission (Cummings and Zoghbi, 2000a; Cummings and Zoghbi, 2000b).

#### **1.2.2.1. Fragile X syndrome (FRAXA)**

FRAXA is an X-linked disorder, which typically presents in males, with mental retardation being the most common feature of the disease, but it is also associated with testicular abnormalities (Jin and Warren, 2000; Kooy *et al.*, 2000). The *FRAXA* locus contains a polymorphic CGG•CCG repeat in the 5' UTR of the *fragile X mental retardation-1 (FMR1)* gene (Fu *et al.*, 1991). The *FMR1* gene product, FMRP, is a selective RNA-binding protein, which is widely, but not ubiquitously, expressed, with particularly high expression levels in neurons and gonads. Expansion of the CGG•CCG repeat beyond 230 trinucleotides causes disease and hypermethylation of the repeat tract, as well as the CpG island within the *FMR1* promoter region (Hansen *et al.*, 1992; Sutcliffe *et al.*, 1992). Promoter hypermethylation results in transcriptional silencing, through histone deacetylation (Chiurazzi *et al.*, 1999; Coffee *et al.*, 1999), leading to low levels of FMRP.

Identification of other mutations in the *FMRI* gene, such as deletions and point mutations among patients showing the usual FRAXA phenotype, but without fragile site expression, firmly established that *FMRI* is the only gene involved in the pathogenesis of FRAXA, and confirmed that the disease results from a loss of FMRP function (Jin and Warren, 2000; Kooy *et al.*, 2000). Given that FMRP shuttles between the nucleus and the cytoplasm, and is associated with actively translating ribosomes in neurons, the phenotypic manifestations associated with FRAXA may result from the potential pleiotropic effects caused by abnormal brain RNA metabolism, essential for synaptic development or maintenance (Jin and Warren, 2000; Kooy *et al.*, 2000).

#### 1.2.2.2. Fragile XE mental retardation (FRAXE)

The FRAXE chromosomal fragile site is associated with a mild form of mental retardation. Similar to FRAXA, FRAXE is caused by an expansion of a polymorphic GCC•GGC trinucleotide repeat immediately adjacent to a CpG island, in the promoter region of a gene termed *FMR2* (Knight *et al.*, 1993). The expanded repeats are also abnormally hypermethylated, leading to transcriptional silencing of *FMR2* (Gecz *et al.*, 1997). The cognitive and behavioural deficits in FRAXE likely result from the transcriptional silencing of the *FMR2* gene and subsequent loss of FMR2 protein function (Cummings and Zoghbi, 2000b). Although the precise function of the *FMR2* gene product is still unknown, its putative role as a transcriptional activator and its high level of expression in areas of the brain involved in learning, memory and emotion, suggests that the pathogenesis in FRAXE results from alterations in neuronal gene regulation (Cummings and Zoghbi, 2000b). An additional gene, *FMR3*, has also been identified in the same region and it shares the same methylated CpG island with *FMR2* (Gecz, 2000). Expression of *FMR3* is also silenced by FRAXE full mutation, and therefore may also contribute to the disease phenotype.

#### 1.2.2.3. Friedreich ataxia (FRDA)

FRDA is a progressive neurodegenerative disorder, involving both the central and peripheral nervous systems. It is autosomal recessive and characterised by progressive ataxia, hypertrophic cardiomyopathy, optic atrophy and diabetes (Timchenko and Caskey, 1999).

FRDA is primarily caused by a large intronic GAA•TTC repeat expansion located in the centre of an *Alu* repeat in the *X25* gene, also known as *frataxin* (Campuzano *et al.*, 1996). Reduced *X25* mRNA results in decreased frataxin protein levels, suggesting that the clinical symptoms result from a partial loss of frataxin function (Bidichandani *et al.*, 1998; Campuzano *et al.*, 1996; Ohshima *et al.*, 1998). In support of this view, patients who carry only one expanded allele and a point mutation within the coding region of the second allele have all the clinical features of typical FRDA (Patel and Isaya, 2001; Puccio and Koenig, 2000). Reduced *X25* expression, maybe the result of transcriptional interference, via the self-association of the GAA•TTC tract, which stabilises a novel higher-order triplex non-B structure, known as “sticky DNA”, with the third DNA strand occupying the major groove of the duplex DNA (Sakamoto *et al.*, 1999). The implications of the FRDA expansion on reduced gene expression is supported by the observation

that GAA•TCC repeats from intron 1 of the *X25* gene inhibit transcription in mammalian cultured cells (Ohshima *et al.*, 1998). Furthermore, intrinsic biochemical properties of GAA•TTC tracts cause reduced transcription elongation *in vitro*, in the absence of any other *frataxin* gene sequences (Grabczyk and Usdin, 2000).

Frataxin has been found to localise in mitochondria, in agreement with the presence of a mitochondrial targeting signal in its amino acid sequence. The protein is likely to be involved in the respiratory function and in iron homeostasis. Frataxin insufficiency is therefore associated with abnormal iron-sulphur homeostasis leading to mitochondria dysfunction, free radical production, oxidative stress and cellular degeneration (Patel and Isaya, 2001; Puccio and Koenig, 2000). The enhanced sensitivity of FRDA fibroblasts to iron and oxidative stress provides further support to this hypothesis (Wong *et al.*, 1999). Expression studies in mice and humans showed that *X25* gene expression is higher in tissues with greater mitochondrial content, and is broadly correlated with the primary sites of FRDA pathology (Patel and Isaya, 2001; Puccio and Koenig, 2000).

#### 1.2.2.4. Myotonic dystrophy type 1 (DM1)

Myotonic dystrophy or *dystrophia myotonica* (DM) type 1 is an autosomal dominant multisystemic disorder, with highly variable manifestations and clinical anticipation. The classical adult onset form of DM1 is primarily characterised by myotonia (delayed muscle relaxation), muscle weakness, and progressive muscle wasting. Other features may include facial dysmorphism, presenile cataracts, testicular atrophy, premature balding in males, kidney failure, hyperinsulin secretion and cardiac muscle conduction abnormalities. Atrophy of facial muscle produces a characteristic haggard appearance. Mental retardation, as well as swallowing and speech difficulties, is sometimes observed. There is also variable loss of mental function, but this is more common in congenital DM1, which is the most severe form of the disorder, also associated with hypotonia, respiratory distress at birth and development abnormalities (Harper, 1998).

DM1 is caused by an expanded CTG•CAG trinucleotide repeat tract in the 3' UTR of the *dystrophia myotonica protein kinase* (*DMPK*) gene (Aslanidis *et al.*, 1992; Brook *et al.*, 1992; Buxton *et al.*, 1992; Fu *et al.*, 1992; Harley *et al.*, 1992; Mahadevan *et al.*, 1992). The gene is expressed predominantly in smooth, skeletal and heart muscle, and at low levels in brain and endocrine tissues (Jansen *et al.*, 1992b).

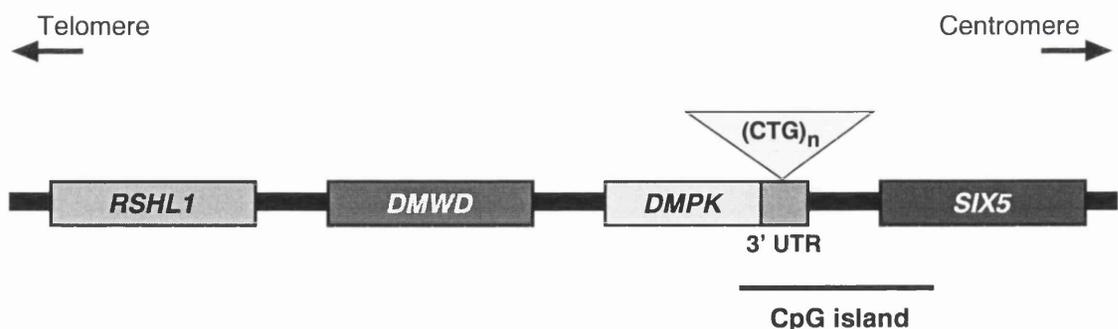
The underlying molecular mechanism by which the expanded CTG•CAG repeat causes the DM1 phenotype is not fully understood, but some critical clues have recently emerged. DM1 is likely to have a complex pathophysiology and a number of mechanisms may contribute, either mutually or exclusively to the disease: haploinsufficiency of DMPK protein, local chromatin effects on the expression of neighbouring genes and novel gain-of-functions conferred on the expanded *DMPK* mRNA (Cummings and Zoghbi, 2000a; Cummings and Zoghbi, 2000b; Richards, 2001).

DMPK is predicted to have several functions based on putative substrates and interacting proteins, including the modulation of skeletal muscle sodium channels (Mounsey *et al.*, 2000),

calcium homeostasis (Benders *et al.*, 1997), RNA metabolism and cell stress response (Timchenko and Caskey, 1999). The CTG•CAG repeat could indirectly alter DMPK protein levels by interfering with gene transcription, RNA processing and/or translation (Fu *et al.*, 1993), resulting in abnormal phosphorylation of downstream substrates. The effect of the CTG•CAG expansion on DMPK levels and the protein role in disease manifestation has been, however, controversial. Although in a majority of patients DMPK protein levels were reduced (Fu *et al.*, 1993), some cases of DMPK activation have also been described (Sabouri *et al.*, 1993). Contradictory reports on the levels of *DMPK* RNA levels detected in DM1 patients and cell lines were also contradictory (Bhagwati *et al.*, 1996; Carango *et al.*, 1993; Hofmann-Radvanyi *et al.*, 1993). A possible explanation for the different results may result from the disrupted processing of *DMPK* transcripts in DM1. Indeed, it was demonstrated that DM1 patients show abnormalities with *DMPK* RNA polyadenylation, as polyA<sup>+</sup> *DMPK* mRNA levels were reduced more significantly compared with total RNA (Hamshere *et al.*, 1997; Krahe *et al.*, 1995; Wang *et al.*, 1995). In addition, abnormal accumulation of *DMPK* transcripts into stable nuclear foci was found in DM1 patients (Davis *et al.*, 1997; Hamshere *et al.*, 1997; Taneja *et al.*, 1995). The nuclear retention of expansion-derived transcripts may not only explain a loss-of-function mechanism, but may also support a gain-of-function for the nuclear-retained transcripts. Unlike FRDA and FRAXA, no confirmed cases of DM1 have been attributed to mutations in the *DMPK* gene, other than the expansion of the CTG•CAG repeat. Additionally, patients who are homozygous for the DM1 mutation do not differ in clinical expression from typical DM1 heterozygotes (Cobo *et al.*, 1993). Moreover, mice lacking *DMPK* develop minor myopathies and muscle weakness (Jansen *et al.*, 1996; Reddy *et al.*, 1996), whereas those overexpressing *DMPK* develop hypertrophic cardiomyopathy and neonatal lethality (Jansen *et al.*, 1996), suggesting that changes in *DMPK* expression levels alone are not sufficient to give rise to the full spectrum of the DM1 phenotype.

A more general effect of the mutation on chromatin structure is suggested by the fact that CTG•CAG repeats are strongly associated with histones (Godde and Wolffe, 1996; Wang *et al.*, 1994; Wang and Griffith, 1995). It is possible that the accessibility of nuclear proteins, such as transcription factors, to the *DM1* locus could affect gene expression. The identification of two other genes flanking *DMPK* raised the possibility that the expanded repeat might alter chromatin structure and alter neighbouring gene expression (Figure 1.2). Sequence analysis of the repeat region showed that the CTG•CAG tract is located within the promoter region of a downstream homeobox gene, initially termed *dystrophia myotonica-associated homeodomain protein* (*DMAHP*), now known as *sine oculis related homeobox 5* (*SIX5*) (Boucher *et al.*, 1995). As with variable *DMPK* levels in DM1, contradictory observations on *SIX5* expression were reported. It was initially shown that no gross change in the level of expression of *SIX5* was observed in DM1 fibroblast cell lines (Hamshere *et al.*, 1997). However, other studies indicated that there could be an allele-specific CTG•CAG repeat length effect on the expression levels of *SIX5* in muscle biopsy specimens, myoblasts and myocardium (Klesert *et al.*, 1997; Korade-Mirnic *et al.*, 1999; Thornton *et al.*, 1997). Heterozygous deletion of *Six5* is sufficient to cause ocular cataracts in transgenic mice (Klesert *et al.*, 2000; Sarkar *et al.*, 2000), and therefore *SIX5* haploinsufficiency, brought about by

the CTG•CAG expansion, is likely to contribute to the DM1 eye phenotype. Indeed, the expression pattern of *SIX5* in the normal adult eye matches the sites of the ocular pathology in DM1, while *DMPK* is not expressed in adult lens (Winchester *et al.*, 1999). Since *SIX5* protein binds to the murine *sodium and potassium ATPase  $\alpha 1$  subunit gene (*Atpa1*)* (Harris *et al.*, 2000), it is not unreasonable to suggest that *SIX5* dysfunction may cause cataracts by affecting ion balance in the lens. The third gene in the *DM1* locus is the *dystrophia myotonica-containing WD repeat motif (DMWD)*, located immediately upstream to *DMPK* (Figure 1.2). The murine homologue is highly expressed in brain, testis, heart and kidney, but absent in skeletal muscle. The tissue-specific expression profile of *DMWD* makes this gene another candidate, whose loss-of-function could contribute to some clinical features of DM1, particularly mental retardation, testicular atrophy and kidney failure (Jansen *et al.*, 1995). Conflicting data also exist regarding the expression levels of *DMWD* in DM1 patient material, with no changes in *DMWD* mRNA cytoplasmic levels (Hamshere *et al.*, 1997), dramatic allele-specific effects (Alwazzan *et al.*, 1999) and an inverse relationship between the size of the CTG•CAG expansion and levels of cytoplasmic polyA<sup>+</sup> mRNA from *DMWD* gene in muscle biopsies of DM1 patients being reported (Eriksson *et al.*, 1999). A fourth gene was recently brought into the picture. A mammalian homologue of a *radial spokehead-like gene (RSHL1)* maps at the *DM1* locus (Figure 1.2), upstream to the *DMWD* gene (Eriksson *et al.*, 2001). Given the importance of radial spokehead proteins in ciliary or flagellar action in the lower organisms, and the expression of the human *RSHL1* in the adult testis, it is not unreasonable to speculate that altered *RSHL1* expression levels in DM1 patients could account for male infertility, characteristic of this condition (Eriksson *et al.*, 2001).



**Figure 1.2. Genomic organisation of the human *DM1* locus.**

The figure shows the four genes mapped within the *DM1* locus on human chromosome 19q13.3. The CTG•CAG repeat expansion is located within the 3' UTR of the *DMPK* gene. The CpG island extends from the last intron of the *DMPK* gene to the first intron of *SIX5*. The promoter region of *SIX5* is situated within the CpG island.

Another hypothesis to explain the variety of clinical features seen in DM1 proposes a gain-of-function model at the RNA level. It is based on a *trans*-dominant effect of the CUG-expanded transcript, which interferes with the normal processing and/or metabolism of numerous RNAs

(Morrone *et al.*, 1997; Taneja *et al.*, 1995; Wang *et al.*, 1995). Expansion of CUG repeats could create RNA binding sites for specific CUG-binding proteins. Overexpression of these RNA binding sites would cause sequestration of specific proteins, resulting in a general effect on the cell RNA metabolism. This hypothesis is supported by the identification of a CUG-binding protein (CUG-BP), a regulator of pre-mRNA splicing (Timchenko *et al.*, 1996a; Timchenko *et al.*, 1996b). Although it was initially hypothesised that CUG-BP levels should be reduced due to sequestration by CUG repeats, CUG-BP levels and activity are in fact enhanced in DM1 cells (Timchenko *et al.*, 1996a), partially through an increased half-life following titration of CUG-BP from a free pool into RNA-protein complexes, through direct interaction with CUG expanded repeats (Timchenko *et al.*, 2001a). It was also reported that CUG-BP might be a substrate for DMPK, and that decreased levels of DMPK in DM1 patients could cause hypophosphorylation and nuclear retention of CUG-BP, thereby establishing an autoregulatory loop (Roberts *et al.*, 1997). Activation of nuclear CUG-BP activity in DM1 results in alteration of RNA processing of CUG-BP dependent RNAs, being therefore considered a key mediator of a *trans*-dominant effect of the DM1 mutation. Aberrant RNA splicing of human cardiac troponin T (cTNT) (Philips *et al.*, 1998), insulin receptor (Savkur *et al.*, 2001) and tau protein (Sargent *et al.*, 2001) have all been reported and associated with increased levels of CUG-BP activity in DM1 cells. In addition, CUG-BP is involved in the translation of a transcription factor, CCAAT/enhancer binding protein  $\beta$ , (C/EBP $\beta$ ) which plays an important role in cell proliferation and differentiation (Timchenko *et al.*, 1999). Indeed, translocation of CUG-BP from the cytoplasm into the nucleus, through the association with expanded CUG repeats, leads to a significant reduction of proteins responsible for cell cycle control (Timchenko *et al.*, 2001b). Furthermore, expanded DM1 CUG repeats bind and activate the double-stranded RNA-activated protein kinase (PKR) in a length-dependent manner (Tian *et al.*, 2000). PKR is involved in regulating cell proliferation and stress responses in mammalian cells (Tian *et al.*, 2000; Williams, 1999). It now seems very likely that a family of proteins that bind CUG repeats, as well as their dependent RNAs, might be affected by the CTG•CAG expansion at the *DM1* locus, mediating the multisystemic character of this disorder.

CUG repeats form unusual and stable RNA hairpins in their natural sequence context of the *DMPK* gene transcript (Michalowski *et al.*, 1999; Napierala and Krzyosiak, 1997). Theoretically, large hairpins may sterically block RNA export through nuclear pores accounting for the nuclear accumulation of *DMPK* transcripts in the nucleus (Koch and Leffert, 1998). Most importantly, the entire hairpin stem forms a perfect double-stranded RNA structure that serves as a binding site for triplet repeat expansion (EXP) proteins (Miller *et al.*, 2000), while CUG-BP localises to the base of the RNA hairpin (Michalowski *et al.*, 1999). The EXP proteins are homologous to the *Drosophila* muscleblind proteins, required for terminal differentiation of muscle and photoreceptor cells (Artero *et al.*, 1998; Begemann *et al.*, 1997). Rather than a single human muscleblind gene, three homologues have been identified in humans: *MBNL*, *MBLL* and *MBXL* (Fardaei *et al.*, 2002), and the three human muscleblind related proteins co-localise with nuclear foci in DM1 (Fardaei *et al.*, 2001; Fardaei *et al.*, 2002; Mankodi *et al.*, 2001).

Crucial experimental support for the toxic RNA effect in DM1, came from both cell culture systems and transgenic mice. Overexpression of long CTG repeats in myoblast culture cells inhibits myogenesis (Amack *et al.*, 1999; Bhagwati *et al.*, 1999; Sabourin *et al.*, 1997), whereas transgenic mice expressing large untranslated CUG repeats inserted into an unrelated mRNA, under the control of the *human skeletal actin* promoter, developed myotonia and myopathy, with no evidence of muscle wasting (Mankodi *et al.*, 2000). In addition, transgenic mice carrying the human genomic DM1 region, which includes *DMWD*, *DMPK* and *SIX5*, with a large CTG•CAG expansion (at least 300 repeat units), display abnormalities in both skeletal muscle and brain, consistent with those observed in DM1 patients (Seznec *et al.*, 2001). These findings imply that expanded CUG repeats are indeed sufficient to generate at least some aspects of the DM1 phenotype.

Defining a second human mutation that causes the multisystemic clinical phenotype of DM1 provided vital support for the pathogenic RNA hypothesis. Myotonic dystrophy type 2 (DM2), which shows remarkable clinical similarity to DM1, is caused by a large untranslated CCTG tetranucleotide expansion within intron 1 of the *zinc finger protein 9 (ZNF9)* gene on chromosome 3, in a locus where genes homologous to *DMPK*, *SIX5* and *DMWD* are absent (Liquori *et al.*, 2001). The common theme resides in the accumulation of *ZNF9* transcripts in discrete nuclear foci (Liquori *et al.*, 2001), which also bind *MBNL*, *MBLL* and *MBXL* (Fardaei *et al.*, 2002; Mankodi *et al.*, 2001). These RNA foci are likely to cause global disruptions in RNA splicing and cellular metabolism, and may well be the unifying mechanism, which underlies all forms of myotonic dystrophy.

#### 1.2.2.5. Spinocerebellar ataxia type 8 (SCA8)

SCA8 is the result of the expansion of an untranslated CTG•CAG repeat. Although the repeat is not translated, it is present in the 3' terminal exon of a processed transcript (Koob *et al.*, 1999). A second brain-specific mRNA was isolated in an orientation opposite to that of the *SCA8* transcript, and found to encode a protein termed kelch-like 1 (*KLHL1*), but this transcript does not include the repetitive sequence (Nemes *et al.*, 2000). Nonetheless, it is possible that the *SCA8* transcript may regulate the expression of *KLHL1*. The pathology of *SCA8* is intriguing, as there is a CTG•CAG repeat size range that is pathogenic, with shorter and larger repeats not resulting in disease. Although this phenomenon is still not yet understood, it is possible that these very large repeats interfere with *SCA8* expression or confer altered RNA processing and/or stability, so that the toxic gain-of-function does not happen (Cummings and Zoghbi, 2000b). Alternatively, *SCA8* transcripts containing alleles in the pathological range can be exported to the cytoplasm, where they exert their effect on the *KLHL1* transcript, while transcripts with larger expansions are retained in the nucleus (Usdin and Graczyk, 2000). It is interesting to note that although the *SCA8* mutation shares molecular similarities to the DM1 mutation, *SCA8* does not exhibit the multisystemic features of DM1. The divergence between these two disorders may be accounted for by the cerebellum-specific expression pattern of the *SCA8* transcript (Koob *et al.*, 1999).

### 1.2.2.6. Spinocerebellar ataxia type 12 (SCA12)

SCA12 is a rare disease caused by a non-coding CAG•CTG trinucleotide repeat expansion in the 5' UTR of the brain-specific *protein phosphatase 2, regulatory subunit  $\beta$*  (*PPP2R2B*) gene (Holmes *et al.*, 1999). Although the repeat is flanked by transcriptional start sites and conserved promoter elements, it is not known whether the expanded repeat is associated with transcriptional interference of the *PPP2R2B* gene or not.

## 1.3. Repeat dynamics

Trinucleotide repeat tracts are intrinsically unstable DNA sequences, as initially revealed by diffused hybridisation signals observed on Southern blot analysis of restriction fragments of genomic DNA containing expanded trinucleotide repeats. The heterogeneous smears were later shown to comprise multiple unresolved bands, each one derived from a single allele carrying a specific repeat number (Monckton *et al.*, 1995; Wörhle *et al.*, 1995). Expanded trinucleotide sequences can indeed undergo changes in repeat number. Although both expansions and contractions have been detected, there is a massive bias towards expansion, not only in the germline but also in somatic cells. The expansion mutation is strongly dependent upon the repeat tract length, such that the probability of expansion increases with the number of tandem repeats, hence the term “dynamic mutations” (Richards and Sutherland, 1992). The molecular bases of repeat instability are not well understood, nor is it known whether somatic and germline instabilities share common mechanisms of expansion.

### 1.3.1. Germline instability

Clinical anticipation refers to the increased severity of the phenotype and earlier age of onset in successive generations of an affected family, and has been described as a hallmark of trinucleotide repeat disorders. Anticipation has now been explained at the molecular level by the positive correlation of disease severity and the inverse correlation of the age of onset with the inherited repeat length, and the propensity of the repeat to increase in length when transmitted through the germline (Harper *et al.*, 1992). Repeats within the non-disease range are relatively stable when transmitted from one generation to the next, and are believed to only rarely expand in steps of one or a few repeats. However, with increasing size of the repeat the chance of further expansion rises dramatically, and disease-associated alleles typically show large changes in size on transmission from parent to offspring (Cummings and Zoghbi, 2000b; Richards, 2001).

The parental origin of the disease allele can influence anticipation. Some expansions are predominantly paternally inherited, while others show a maternal bias. For most of these disorders, especially those involving smaller repeat expansions (~35-80 repeats), such as polyglutamine diseases, there is greater risk of repeat expansion upon paternal transmission. In contrast, FRAXA and FRDA, which involve very large expansions (~200-2000 repeats), are most often associated

with maternal transmission (Jin and Warren, 2000; Kooy *et al.*, 2000; Patel and Isaya, 2001). DM1 presents rather more complex sex-related differences in the intergenerational amplification of the CTG•CAG repeat. Small expansions (from 37, up to ~200 repeats) are most unstable when transmitted by males, giving rise to the excess of transmitting grandfathers in DM1 pedigrees (Brunner *et al.*, 1993), whilst the largest expansions (up to several thousand repeats), associated with congenital DM1, are usually transmitted by females (Lavedan *et al.*, 1993; Tsilfidis *et al.*, 1992). Whether sex-of-origin effects reflect some sex-specific differences in the mutational pathway, or some sort of selection, either in the spermatogenic cells or in the sperm cells directly, remains unclear. The analysis of the grossly unstable SCA7 alleles shed some light on this subject. Single molecule analysis of sperm DNA revealed that the vast majority of the mutant sperm cells of two SCA7 males contained alleles that were so large that most of the affected offspring would, at best, have a severe infantile form of the disease in the next generation. The under-representation of such very large expanded alleles in patients suggests that a significant proportion of such alleles might be associated with dysfunctional sperm or embryonic lethality (Monckton *et al.*, 1999). In an interesting exception, sperm analysis of 500-800 CTG•CAG repeat tracts at the SCA8 locus, displayed strong bias towards contraction (Moseley *et al.*, 2000).

The parent-of-origin effect has been recreated in some animal mouse models carrying transgenic trinucleotide repeats. Nevertheless, repeat number changes through mouse germline transmission, are considerably smaller than in humans, involving only a few CAG•CTG repeats. Knock-in HD mouse models exhibit modest germline instability ( $\pm 1-5$  repeat units), showing an expansion bias in male transmission, and a deletion bias upon maternal transmission (Shelbourne *et al.*, 1999; Wheeler *et al.*, 1999b). A knock-in transgenic mouse model of FRAXA, in which a long (CGG•CCG)<sub>98</sub> repeat was inserted into the murine *Fmr1* gene by homologous recombination, showed moderate expansion-biased triplet repeat instability upon both maternal and paternal transmission, with a tendency for longer repeat changes (both expansions and deletions) to occur during paternal transmission (Bontekoe *et al.*, 2001).

Pedigree analyses have attempted to estimate human germline instability from the relationship of the allele sizes in the parent and their offspring by direct comparison of the repeat numbers in their somatic cells, but only one report confirmed that a CAG•CTG repeat expands with increasing age of the father (Sato *et al.*, 1999). Nonetheless, sizing of individual HD sperm cells revealed that both mutation frequency and mean change in allele size increases with increasing somatic repeat number, and suggested a paternal age effect on expansion of the CAG•CTG repeat (Leefflang *et al.*, 1999). In addition, pedigree data of trinucleotide repeat germline instability in mouse models carrying expanded CAG•CTG repeat tracts derived from the human *DM1* and *DRPLA* loci, revealed an age-dependent bias towards expansion with age in male transmissions (Mangiarini *et al.*, 1997; Sato *et al.*, 1999; Seznec *et al.*, 2000; Zhang *et al.*, 2002). Similarly, enhanced deletion-biased germline instability, concerning both the overall mutation frequency and the extent of repeat contraction, has been associated with advanced age of the transmitting mother in SCA1 (Kaytor *et al.*, 1997) and DRPLA (Sato *et al.*, 1999) transgenic mice.

Contradictory reports have been published regarding segregation distortion in favour of mutant alleles in some CAG•CTG trinucleotide repeat disorders. Single sperm analyses, performed in Japanese SCA3 patients, revealed a segregation ratio between sperm with an expanded allele and those with a normal allele of ~6:4, significantly different from the expected 1:1 segregation ratio, suggesting meiotic drive in male meiosis (Takiyama *et al.*, 1997). In contrast, the data collected from a similar study performed on French patients, did not support gametic segregation distortion in favour of expanded SCA3 alleles, showing that the segregation ratio of single sperm with an expanded allele was not statistically different from the expected 1:1 segregation ratio (Grewal *et al.*, 1999). Similarly, single sperm analyses did not reveal meiotic drive in DM1 (Leeflang *et al.*, 1996; Monckton *et al.*, 1995), HD (Leeflang *et al.*, 1995) or DRPLA (Takiyama *et al.*, 1999).

In summary, trinucleotide repeats exhibit expansion-biased germline instability, which does not entirely depend on the repeat tract length *per se*, being also dependent on yet unidentified factors, intimately associated with the sex of the transmitting parent.

### **1.3.2. Somatic instability**

Repeat size differences between somatic cells have been documented in most trinucleotide repeat disorders. In DM1 families, the enormous fluctuations in repeat length, both between and within tissues, is outstandingly evident and easy to study. The high prevalence of the disorder, the availability of large families, and tissue biopsy samples render DM1 a particularly interesting model in which to study somatic dynamic mutations. Therefore, in this section, greater focus will be given to the dynamics of expanded CTG•CAG repeats in the soma of DM1 patients, and a parallel with other conditions will be presented when appropriate.

The extensive somatic instability of DM1 CTG•CAG repeats is well documented, with the expanded allele frequently presenting as a smear on restriction digested genomic DNA Southern blot analysis for a wide range of tissues, including peripheral blood lymphocytes, liver, brain and heart (Jansen *et al.*, 1994; Lavedan *et al.*, 1993; Shelbourne *et al.*, 1992). Moreover, genomic Southern blot procedures were able to detect large tissue-to-tissue variability in the repeat size, between skeletal muscle and blood (Anvret *et al.*, 1993; Ashizawa *et al.*, 1993; Lavedan *et al.*, 1993; Thornton *et al.*, 1994). Cloning DM1 cells proved that heterogeneous smears result from somatic mosaicism of cells carrying different repeat lengths (Wörhle *et al.*, 1995). CAG•CTG tissue-specific somatic instability has also been reported in HD (Telenius *et al.*, 1994), SCA1 (Chong *et al.*, 1995; Chung *et al.*, 1993), DRPLA (Takano *et al.*, 1996) and SBMA (Tanaka *et al.*, 1999). Although associated with different trinucleotide repeat expansions, tissue-specific somatic mosaicism has also been described in FRAXA (Wohrle *et al.*, 1996) and FRDA (Bidichandani *et al.*, 1999; Hellenbroich *et al.*, 2001)

Moderate somatic instability becomes apparent during foetal development. Analysis of five DM1 fetuses and two DM1 neonates, using standard Southern blot procedures, revealed that low CTG•CAG repeat size heterogeneity between tissues was detectable between 13 and 16 weeks of gestation (Martorell *et al.*, 1997), with the largest expansions being found in kidney, adrenal gland,

ovary, diaphragm, heart and brain (Wörhle *et al.*, 1995). Consistent with the prenatal development of somatic mosaicism, a congenital DM1 newborn was reported to exhibit somatic repeat instability soon after birth in a wide range of tissues (Wong and Ashizawa, 1997). However, analyses of peripheral blood DNA from adult DM1 patients, revealed that an increase in the average repeat length could be detected over a time span of two to five years (Martorell *et al.*, 1995; Wong *et al.*, 1995), establishing that the accumulation of repeat length mutations is certainly not limited to embryogenesis. Analysis of somatic mosaicism in autopsied brains of DRPLA patients also showed that the degree of somatic mosaicism increases with age (Takano *et al.*, 1996). A more detailed study of DM1 somatic instability, using sensitive single molecule analysis, confirmed the progression of expansion-biased repeat size heterogeneity over time, and established a correlation between the average repeat number, the inherited CTG•CAG repeat length and the time interval (Martorell *et al.*, 1998). It is therefore assumed that in DM1, and possibly in most of trinucleotide repeat diseases, somatic instability proceeds throughout adult life, through a directional pathway, which involves “step-wise” gains of small number of repeat units, resulting in an increasing overall range and mean size, and decreasing modal allele frequency, over time. The rate of progression appears to be proportional to progenitor allele length and dependent on tissue type.

Neither the molecular mechanisms underlying repeat instability, nor the bases of variability between tissues at any of the expanded repeat loci are understood. Further analysis of these processes in patients is limited by the availability of appropriate samples throughout the lifetime of an individual, and is further compromised by inter-individual genetic and environmental variation. Transgenic mouse models have therefore been generated, in order to provide powerful animal systems with which to assess trinucleotide dynamics. Somatic CAG•CTG repeat instability observed in DM1 patients was detected in transgenic mice for a large genomic fragment of the human *DM1* locus (Gourdon *et al.*, 1997; Lia *et al.*, 1998; Seznec *et al.*, 2000) or for the 3' UTR of the *DMPK* gene (Fortune *et al.*, 2000; Monckton *et al.*, 1997), and also in knock-in mouse model of DM1, in which an expanded CTG•CAG repeat was inserted in the endogenous mouse *Dmpk* gene by targeted recombination (van Den Broek *et al.*, 2002). Transgenic and knock-in mice carrying an expanded CAG•CTG repeat derived from a mutated *HD* gene have also recreated somatic repeat instability observed in HD patients (Kennedy and Shelbourne, 2000; Mangiarini *et al.*, 1997; Wheeler *et al.*, 1999b). In addition, transgenic murine lines, harbouring a single copy of a full-length expanded *DRPLA* gene, randomly integrated into their genome, have also exhibited age-dependent, tissue-specific somatic mosaicism (Sato *et al.*, 1999). These findings established that mouse models are capable of recreating large repeat length changes in the soma, similar to those observed in human patients. Transgenic mice have therefore provided a valuable tool to investigate the molecular mechanisms underlying somatic trinucleotide repeat instability in mammalian systems, and created new avenues to assess the complex dynamics of simple repetitive DNA sequences.

### 1.3.3. *Dmt transgenic mice*

The availability of genetically and biochemically defined systems in *E. coli* and *S. cerevisiae* have enabled detailed investigation of the mechanisms that drive trinucleotide repeat instability in these simple organisms. Although major differences between mammalian, bacterial and yeast cells exist in terms of DNA replication enzymology and chromatin structure, studies in bacteria and yeast have provided numerous insights into the factors affecting stability of repeated DNA sequences. However, it is clear that complete comprehension of the biology of the triplet repeat diseases requires studies in more complex systems. To facilitate more detailed studies of repeat dynamics and the mutational mechanisms underlying instability throughout mammalian development, Monckton *et al.* (1997) generated five transgenic mouse lines incorporating expanded CAG•CTG arrays derived from the human *DM1* locus. Transgene analysis has revealed germline instability, exhibiting expansions and deletions, parent-of-origin effects and segregation distortion (Monckton *et al.*, 1997; Zhang *et al.*, 2002). Trinucleotide repeat somatic instability has been extensively studied particularly in two mouse lines, both carrying a single copy of the transgene. *Dmt-D* mice showed age-dependent, tissue-specific somatic mosaicism, whilst the transgene showed remarkable stability in the soma of *Dmt-E* mice (Fortune *et al.*, 2000). An investigation into the integration sites has revealed that the transgene is flanked by DNA sequences of chromosome 11, syntenic to human chromosome 17, in the *Dmt-D* line, while it integrated into a long interspersed nuclear element family 1 (LINE1) in the *Dmt-E* line (G.J. Brock and D.G. Monckton, personal communication). Mutational differences between the lines illustrate that triplet repeat stability in mice, as in humans, is not purely a function of repeat length, but must be also modulated by as yet unidentified *cis*-acting genetic modifiers. On the other hand, tissue-specific patterns of somatic mosaicism in *Dmt-D* mice suggest the involvement of tissue- and/or cell-specific *trans*-acting factors in the control of the dynamics of expanded trinucleotide repeats.

### 1.3.4. *Possible link between somatic instability and pathophysiology*

Somatic instability of trinucleotide repeats involves a dynamic process in which the repeat size increases with age, at different rates in various tissues and therefore it appears logical to assume that trinucleotide repeat somatic mosaicism plays an important role in tissue- and age-specific phenotypic variability, and that it is most certainly correlated with the clinical disease progression. Mounting evidence has indeed provided strong support to this hypothesis.

In DM1 patients, expansion of the CTG•CAG repeat shows a positive correlation with the severity of the disease (Hunter *et al.*, 1992). More interestingly, much larger repeat expansions are consistently observed in the major affected tissue, skeletal muscle, relative to peripheral blood leukocytes from the same individual (Anvret *et al.*, 1993; Ashizawa *et al.*, 1993; Monckton *et al.*, 1995; Thornton *et al.*, 1994). Very little is known about the relative stability of the repeats in additional DM1 tissues, with far fewer observations having been reported. General findings suggest that in addition to skeletal muscle, heart and kidney usually display the largest expansions (Jansen

*et al.*, 1994; Joseph *et al.*, 1997; Lavedan *et al.*, 1993; Tachi *et al.*, 1995; Tachi *et al.*, 1993; Wörhle *et al.*, 1995), possibly associated with heart conduction defects and kidney failure exhibited by DM1 patients. In addition, longer average repeat lengths and broader ranges of variability have been found in older patients, usually presenting a more severe DM1 clinical picture (Martorell *et al.*, 1998; Monckton *et al.*, 1995; Wong *et al.*, 1995).

An inverse correlation has been established between the amount of frataxin protein and the expansion size of the smaller *FRDA* allele in tissues from human patients (Campuzano *et al.*, 1997). Limited somatic mosaicism has been described in *FRDA* blood leukocytes by low DNA input polymerase chain reaction (PCR) amplification across the GAA•TTC repeat (Bidichandani *et al.*, 1999; Hellenbroich *et al.*, 2001). If trinucleotide repeat instability is extended to other somatic tissues, it could be speculated that somatic cells carrying longer alleles would be more dramatically affected, exhibiting more severe defects in mitochondrial metabolism and higher sensitivity to oxidative stress. Although speculative, such a scenario could account for the phenotypic variability observed in this condition.

*FRAXA* patients carrying fully expanded CGG•CCG repeats (>250 units) exhibit *FMRI* promoter methylation, associated with transcriptional silencing and lack of FMRP. However, subtle fragile X-like features have been described in premutation carriers (~60-250 repeat units) (O'Donnell and Warren, 2002). Although controversial, positive correlations between the intermediate repeat size, cognitive impairment (as assessed by intelligence quotient tests) (Kaufmann *et al.*, 1999; Tassone *et al.*, 1999), interpersonal sensitivity and depression (Johnston *et al.*, 2001) and ovarian failure (premature menopause) (Allingham-Hawkins *et al.*, 1999) have been reported in premutation carriers. In addition intermediate *FRAXA* allele sizes has been shown to correlate negatively with FMRP levels, and positively with *FMRI* transcription (Kenneson *et al.*, 2001), possibly accounting for a parallel between the premutation repeat length and the severity of symptoms. It is not unreasonable to hypothesise that such a parallel might be mediated by somatic expansion in a subset of cells.

Crucial evidence in support of an association between tissue-specific somatic mosaicism and polyglutamine pathology emerged from studies in HD mouse models. The *HD* gene is widely expressed in many tissues both within and outside the central nervous system (Li *et al.*, 1993; Strong *et al.*, 1993). Given that mutant huntingtin is present at similar levels in many central nervous system and non-central nervous system tissues (Li *et al.*, 1993; Strong *et al.*, 1993), it is unclear why medium spiny striatal neurons are selectively vulnerable to the disease process (Graveland *et al.*, 1985). If the degree of CAG•CTG expansions is directly related to the pathogenesis, is it theoretically possible that the neuropathological findings in HD reflect somatic mosaicism with different levels of CAG•CTG expansion in different tissues. Indeed CAG•CTG repeat tracts display tissue-specific mosaicism in autopsied brains of HD patients. The highest instability was found in those regions showing greatest neuropathological involvement and was seen most dramatically in juvenile patients (Telenius *et al.*, 1994). However, the repeat length changes observed in these studies were relatively small, most probably because the methods used to assess repeat size variation were based upon bulk DNA analysis. Using much more sensitive

small pool polymerase chain reaction (SP-PCR) techniques (Monckton *et al.*, 1995), dramatic age-dependent instability of the expanded CAG repeat tract was reported in a knock-in mouse model of HD (Kennedy and Shelbourne, 2000). Most importantly, striking large expansions were restricted to the affected tissue, the striatum, in the absence of progressive striatal loss (Kennedy and Shelbourne, 2000; Shelbourne *et al.*, 1999), commonly observed in post-mortem tissue from end-stage human HD patients. The median mutation length in the striatum was significantly greater than the corresponding measures in both cerebellum and cortex (Kennedy and Shelbourne, 2000). Similarly, very large expansions have also been observed in post-mortem striatum samples of human HD brain material (P.F. Shelbourne, personal communication). These findings clearly suggest that the selective vulnerability of neurons in HD may be caused by cell-specific determinants, which directly or indirectly induce dramatic expansion-biased mutability of the CAG•CTG repeat over time, and provide a crucial clue to the puzzling regional selectivity of neuronal loss in polyglutamine diseases. CAG•CTG repeats, being particularly prone to expansion in the striata of aging HD humans and mice, would render striatal cells more prone to succumb early, in virtue of the fact that they would exceed a critical polyglutamine concentration threshold first.

## 1.4. Modifiers of trinucleotide repeat instability

Mounting evidence suggests that the intrinsic nature of repeating DNA is causative of mutation within trinucleotide repeat sequences. However, *in vitro* studies, human studies, and transgenic mouse models, as well as simple model organisms, have suggested that the process of dynamic mutation is affected by a variety of elements and factors. The nature and degree of instability may depend on the predisposition of a repetitive sequence to form secondary structures, the location of the repeat within the affected gene, cell and tissue type-specific factors, parent-of-origin effects, sequences within and around the repeat, stage of embryonic development, state of epigenetic modification, and possibly other as yet unidentified factors. Such modifiers have been categorised into those directly associated with the expanding repeat (*cis*-acting elements) and those whose interaction with the repeat contributes to its instability (*trans*-acting factors).

### 1.4.1. *Cis*-acting modifiers of trinucleotide repeat instability

Trinucleotide repeat instability is governed by unique *cis*-elements. The chromosomal position may account for interlocus variation in repeat mutability, whereas differences in mutation rates within a single locus (intralocus variation) may be associated with the copy number and the composition of the repeat tract.

The instability of the repeat is sensitive to the length of the repeat tract, as the products of an expansion mutation are more likely to undergo subsequent mutation than the original substrate (Richards and Sutherland, 1992). Unstable trinucleotide repeats are almost always CNG sequences,

whose secondary structure-forming ability is thought to promote instability (Mitas, 1997; Pearson and Sinden, 1998b; Sinden *et al.*, 2002). If expansion were solely driven by repeat length, two alleles of similar size would be expected to exhibit similar levels of instability. That appears not to be the case. Similar sized *HD* alleles, carrying 29-35 repeats, have shown marked differences in expansion rates, as determined by single sperm analysis (Chong *et al.*, 1997). The initial mutation on these alleles is likely a rare event that increases the instability of the repeat, setting in process the subsequent expansion of these few founder mutations (Chong *et al.*, 1997; Goldberg *et al.*, 1995). Male germline mutational pathways in DM1 patients exhibit allele length-independent variation, which does not correlate with progenitor allele size and is clearly distinct from that of the soma (Martorell *et al.*, 2000). These data strongly support a role for *cis*- and *trans*-acting genetic modifiers of DNA metabolism.

Experimental data suggest that the sequences in the vicinity of the repeat may contribute to the mechanism of instability in several disease loci. Haplotype analyses have revealed that specific haplotypes are particularly enriched on disease-associated chromosomes, suggesting that chromosome-specific *cis*-acting factors play a significant role in influencing trinucleotide repeat intergenerational behaviour. Chromosomes with specific flanking haplotypes are associated with expansions in DM1 (Imbert *et al.*, 1993; Mahadevan *et al.*, 1993; Neville *et al.*, 1994), HD (Goldberg *et al.*, 1995), FRAXA (Crawford *et al.*, 2000; Zhong *et al.*, 1996), FRDA (Cossee *et al.*, 1997) and SCA2 (Choudhry *et al.*, 2001). This association may, however, depend on either intra-repeat effects or extra-repeat *cis*-factors. Some disease-associated haplotypes are characterised by loss of interspersed interruptions within the trinucleotide repeat tract (Choudhry *et al.*, 2001; Eichler *et al.*, 1996), supporting the importance of repeat interruptions in maintaining trinucleotide stability. However, some disease-associated haplotypes retained repeat interruptions, hinting that other haplotype-specific *cis*-acting factors, flanking the repeat are important in determining the allele's predisposition to instability and to cause disease. Indeed, a CGG/GGG polymorphism at the 3' end of the SCA3 CAG•CTG repeat affects intergenerational instability (Igarashi *et al.*, 1996). Nevertheless, these factors can only explain a small portion of the variance observed, suggesting that other *cis* and/or *trans*-modifiers may play a role in repeat dynamics.

The influence of immediately flanking sequences has also been proposed as a key factor controlling repeat germline instability at different chromosomal loci. The mutation rate, the size and direction of the length change mutations observed vary, with some loci being apparently more mutable and liable to expand to greater lengths than others (Brock *et al.*, 1999). The instability of an expanded CAG•CTG repeat during germline transmission is indeed correlated to the flanking GC levels: the higher the GC level, the greater the mutation rate of the repeat (Brock *et al.*, 1999). It is of interest that these effects were less pronounced in female transmissions than in male transmissions, suggesting that the mutation process is less susceptible to flanking sequence modifiers in females.

Proximity to a preferred origin of replication may be another major chromosomal component of instability. The distance between the origin of replication and the repeat tract has been described as a key variable that affects repeat stability in a simian virus 40 (SV40) DNA

system replicated within primate COS-1 African green monkey cells (Cleary *et al.*, 2002). A location effect may therefore explain the existence of founder chromosomes, and their possible predisposition to trinucleotide repeat expansion. The *Alu* insertion/deletion polymorphism detected in linkage disequilibrium with expanded *DM1* alleles in the Caucasian population (Mahadevan *et al.*, 1993), might have changed the distance between the origin of replication and the trinucleotide repeat itself, enhancing its propensity to expand and cause disease. In addition, the location of the origin of replication to the repeat tract might vary between tissues, explaining the tissue-specific patterns of somatic mosaicism (Cleary *et al.*, 2002).

The repeat size variability in somatic tissues of different transgenic mouse lines (Fortune *et al.*, 2000; Lia *et al.*, 1998; Mangiarini *et al.*, 1997) has also provided support to the suggestion that an undefined chromosomal component or position might be critical for somatic (Fortune *et al.*, 2000; Lia *et al.*, 1998; Mangiarini *et al.*, 1997) and germline instability (Zhang *et al.*, 2002).

#### **1.4.2. Trans-acting modifiers of trinucleotide repeat dynamics**

While there is indirect evidence for the existence of *trans*-acting modifiers of trinucleotide repeat dynamics (such as gender bias in repeat instability during their transmission from parent to offspring), the identity of such factors remains elusive. *In vitro* studies and data from bacteria, yeast and mouse models, suggest that candidates may include proteins involved in DNA replication and repair (Richards, 2001). DNA polymerase III proofreading activity, for instance, was shown to be essential for maintaining the integrity of long CAG•CTG triplet repeat tracts in *E. coli* (Iyer *et al.*, 2000), suggesting that expansion and deletion events of triplet repeats in bacteria are enhanced by mutations that reduce the fidelity of replication. The proofreading exonuclease activity in *E. coli* probably removes slipped structures, formed in the triplet repeat sequence tracts during replication, thereby significantly reducing the frequency of deletion and expansions. Among the components of the replication machinery, mutations in DNA polymerase  $\delta$ , polymerase  $\epsilon$  and in the proliferating-cell nuclear antigen (PCNA) also destabilise CAG•CTG repeat tracts in yeast (Schweitzer and Livingston, 1999).

In R6/1 mice, transgenic for a single integrated copy of a CAG•CTG repeat in exon 1 of the human *HD* gene (Mangiarini *et al.*, 1996), expansion and contraction germline events appear to be influenced post-zygotically by the gender of the embryo (Kovtun *et al.*, 2000), suggesting that there might be X- and Y-encoded *trans*-acting factors that influence the dynamics of expanded trinucleotide repeats. In contrast, no such effect was identified in *Dmt* mice, transgenic for the 3' UTR of the *DMPK* gene (Zhang *et al.*, 2002).

Regardless of their still speculative nature, it is not unreasonable to imagine that individual-specific factors may also be key determinants of trinucleotide repeat dynamics in humans. The nature of such modifiers might be envisaged to include factors such as DNA repair gene variants or environmental genotoxic agents.

In the next section an overview of the different DNA repair systems is presented, as deficiencies in the DNA repair machinery have been implicated in several human diseases, often associated with genomic instability.

## 1.5. DNA repair systems

DNA repair pathways must operate in all living cells to respond to the endogenous threats to DNA, as well as to toxic external damaging agents. The different mechanisms of DNA repair have proved to be unexpectedly complex and diverse, and have also been implied in surveillance and maintenance of genomic stability.

### 1.5.1. Base excision repair

Base excision repair (BER) is the process used for removing many types of damage produced spontaneously in DNA, such as oxidation or alkylation products. The initial step is the removal of the damaged base by one of a series of lesion-specific DNA glycosylases to generate an abasic (AP) site, either apurinic or apyrimidinic. An incision is made 5' to the AP site by an AP endonuclease. The deoxyribose phosphate at the 5' end of the incised strand is removed by the phosphodiesterase activity of DNA polymerase  $\beta$ , and the resulting one-nucleotide gap is filled in by DNA polymerase  $\beta$  and sealed by either DNA ligase I or DNA ligase III, with the involvement of other accessory proteins (Lehmann, 1998; Moses, 2001).

### 1.5.2. Nucleotide excision repair

The nucleotide excision repair (NER) is an important DNA repair pathway involved in the removal of a wide variety of DNA lesions, including photoproducts induced by ultraviolet (UV) light and chemically-induced bulky lesions, which result in major distortions of the double helical structure. Mammalian NER has been studied extensively in rodent and human fibroblasts as well as in established cell lines. This process involves the products of some 30 genes (Wood, 1997). Genetic defects in individual NER proteins can cause severe human diseases, including Cockayne's syndrome (CS) and xeroderma pigmentosum (XP). Cockayne's syndrome patients manifest extreme dwarfism, severe mental retardation, characteristic retinal pigmentation, photosensitivity and deafness. Xeroderma pigmentosum is associated with high sensitivity to UV light, predisposition to skin cancer and neurological abnormalities (Moses, 2001). In humans, genes required for NER have been identified as the seven repair-deficient complementation groups (groups A to G) of this disease (Cleaver *et al.*, 2001; de Laat *et al.*, 1999; Wood, 1997). The NER mechanism involves large protein complexes, which recognise structural changes caused by DNA damage, and proceeds by a three-step mechanism: dual incisions on both sides of the lesion;

excision of the damaged base in an oligonucleotide approximately 24-31 nucleotide long; filling in of the post-excision gap and ligation (Cleaver *et al.*, 2001; de Laat *et al.*, 1999; Wood, 1997).

NER can be divided into two subpathways. Transcription-coupled repair (TCR) refers to the preferential repair of transcribed strands in active genes (Mellon *et al.*, 1987), while global genome repair (GGR) refers to removal of DNA damage indiscriminate of the transcriptional status of the DNA (Cleaver *et al.*, 2001; de Laat *et al.*, 1999).

In *E. coli*, the deficiency of some NER functions dramatically affects the stability of long CTG•CAG tracts cloned into bacterial plasmids (Parniewski *et al.*, 1999).

### 1.5.3. DNA mismatch repair

The preservation of genomic integrity requires the proper functioning of multiple replication, repair, and recombination processes. DNA mismatch repair (MMR) is one of such processes. The primary function of this system is to eliminate base-base mismatches and insertion-deletion loops, which arise as a consequence of DNA polymerase errors generated during DNA replication. The former lesions typically affect non-repetitive DNA, and lead to single base substitutions. Insertion-deletion loops are usually found within repetitive DNA sequences and are responsible for gains or losses of short repeat units within microsatellites, a phenomenon referred to as microsatellite instability (Buermeyer *et al.*, 1999; Harfe and Jinks-Robertson, 2000). It is therefore conceivable that MMR mutations might somehow affect the stability of trinucleotide repetitive sequences associated with human diseases (for further details on the effects of MMR mutations on the stability of triplet repeats see Chapter 8).

The isolation of *E. coli* strains with elevated frequencies of spontaneous mutations, so-called mutator phenotypes, contributed to the identification of four “mutator” genes that play central roles in MMR: *mutS*, *mutL*, *mutH* and *mutU*. MutS is an ATPase that acts as a homodimer to initially recognise base/base mismatches or small insertion/deletion loops, which arise during DNA replication and escape proofreading by the replicating polymerase. The MutS protein bound to a base mispair is then recognised by the MutL protein. The interaction with MutS results in the activation of a third protein MutH. MutH is an endonuclease that introduces a nick into hemimethylated DNA on the nascent strand. MutL is also required to load MutU at the site of the MutH-induced nick, facilitating DNA unwinding and subsequent exonucleolytic removal of the nascent strand. Finally, DNA polymerase II and DNA ligase are necessary for resynthesis and ligation. The combined action of all proteins directs the repair of mismatches to the newly synthesised DNA strand. This system works in a bidirectional way, and mismatches of up to four bases are efficiently repaired (Buermeyer *et al.*, 1999; Harfe and Jinks-Robertson, 2000). In yeast and mammals, the existence of multiple homologues of *mutS* and *mutL* reflects not only conservation of the MMR pathway, but also both specialisation and overlapping functions for MMR gene products (Table 1.3). Given the similarities, most of the actual understanding of the mammalian MMR is derived from studies of the *E. coli* MutHLS mismatch repair system.

**Table 1.3. *mut* homologues in yeast and humans.**

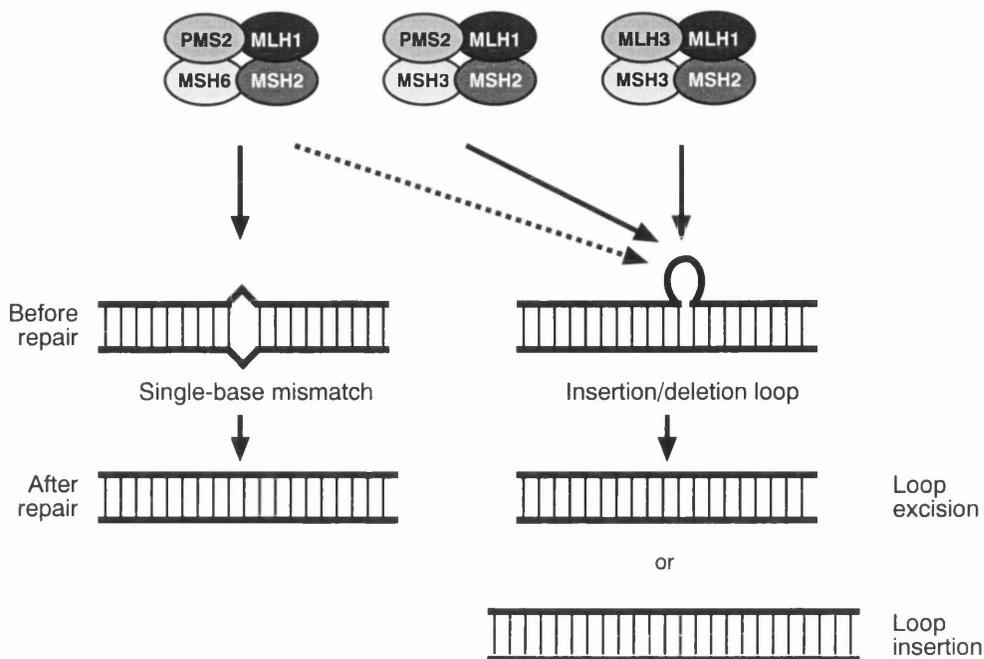
<i>E. coli</i>	<i>S. cerevisiae</i>	<i>H. sapiens</i>
<i>mutS</i>	<i>msh2, msh6, msh3</i>	<i>MSH2<sup>a</sup>, MSH6<sup>a</sup>, MSH3</i>
	<i>msh1</i>	N/I
	<i>msh4, msh5</i>	<i>MSH4, MSH5</i>
<i>mutL</i>	<i>mlh1</i>	<i>MLH1<sup>a</sup></i>
	<i>pms1</i>	<i>PMS2<sup>a</sup></i>
	<i>mlh2, mlh3</i>	<i>PMS1<sup>a</sup></i>
<i>mutH</i>	N/I	N/I
<i>mutU (UvrD)</i>	N/I	N/I

<sup>a</sup> Mutations found in cancer families  
N/I, Not identified

Six MutS homologues have been identified in eukaryotes (MSH1-6). *msh1* is required for normal mitochondrial function in *S. cerevisiae*. Mammalian MSH1 homologues have not been reported to date. Based on studies in yeast and mice, MSH4 and MSH5 are thought to have specialised roles in the control of recombination processes during meiosis and do not appear to function in mutation avoidance. MSH2, MSH3 and MSH6 participate in primary mismatch recognition, with MSH2 functioning as an obligate partner in two heterodimers: MutS $\alpha$ , composed of MSH2 and MSH6; and MutS $\beta$ , composed of MSH2 and MSH3. MutS $\alpha$  primarily binds single-base mismatches as well as single base insertion/deletion mispairs, whereas MutS $\beta$  preferentially recognises 2-4 bp insertion-deletion loops. However, MutS $\beta$  can also bind single base insertion/deletion mispairs (Figure 1.3). The overlapping binding specificities of MutS $\alpha$  and MutS $\beta$  suggest partially redundant roles during MMR (Buermeyer *et al.*, 1999; Jiricny, 2000; Peltomaki, 2001b).

Multiple eukaryotic homologues of MutL have also been identified in yeast and mammalian cells (Table 1.3). As with MutS homologues, the MutL homologues MLH1, PMS2 (*pms1* in yeast) and PMS1 (*mlh3* in yeast) can interact to form two heterodimers: MutL $\alpha$  and MutL $\beta$ . MutL $\alpha$ , which is composed of MLH1 and PMS2, appears to be involved in the repair of both single-base mismatches and insertion/deletion loops. MutL $\beta$  is composed of MLH1 and MLH3, and participates specifically in the repair of insertion/deletion loops (Figure 1.3) (Buermeyer *et al.*, 1999; Harfe and Jinks-Robertson, 2000; Jiricny, 2000).

In eukaryotes, there appears to be some functional overlap between NER and MMR pathways, and certain gene products are required for both processes. Human tumour cell lines defective in MSH2 or MLH1 perform less TCR on UV-induced damage (Leadon and Avrutskaya, 1997). MMR proteins are also involved in TCR of oxidative damage, as MutS $\alpha$ -defective human cells are unable to remove thymine glycol from the transcribed strand of an active gene (Leadon and Avrutskaya, 1997).



**Figure 1.3. Human DNA MMR protein complexes.**

The figure shows different human DNA MMR complexes and their preferences for different types of mismatches. Mismatches are shown before and after correction (adapted from Peltomaki, 2001a).

### 1.5.3.1. MMR, microsatellite instability and cancer

MMR plays an important role in the prevention of tumourigenesis. DNA lesions, if not properly repaired, may be responsible for the development of human cancers. In humans, defective MMR is causative of hereditary non-polyposis colorectal cancer (HNPCC). HNPCC is a common cancer predisposition syndrome characterised by a dominant mode of transmission and high penetrance. HNPCC patients develop predominantly colon, endometrial and ovarian tumours, as well as malignancies of the stomach, pancreas, small intestine, skin, breast and urinary tract (Peltomaki, 2001a). Mutations in the human MMR genes *MSH2* and *MLH1* are a frequent cause of predisposition to HNPCC (Peltomaki, 2001b), with 70% of the germline mutations identified producing truncated forms of the protein, and the remainder being missense mutations (Toft and Arends, 1998). Genomic deletions at the *MSH2* locus are also another frequent cause of HNPCC (Wijnen *et al.*, 1998). In contrast, mutations in *PMS2*, *MSH3*, *MSH6* or *MLH3* are rarely correlated with cancer (Peltomaki, 2001b). In addition to mutations, epigenetic inactivation of MMR genes may also be responsible for microsatellite instability in non-hereditary types of cancer. Such is the case for *MLH1* promoter hypermethylation in the majority of patients suffering from sporadic endometrial carcinoma (Gurin *et al.*, 1999). Loss of DNA MMR is a rate-limiting step in the

aetiology of tumours associated with HNPCC. Affected HNPCC family members inherit a mutant allele in one of the MMR genes. The presence of a wild-type allele appears to be sufficient for normal MMR activity. Cancer progression, in predisposed individuals, results from the loss of the wild-type allele, through somatic mutation, promoter methylation or loss of heterozygosity (Peltomaki, 2001a; Peltomaki, 2001b; Toft and Arends, 1998).

MMR deficiencies can result in increased mutation rates, 100- to 600-fold above normal levels, and therefore lead to an overall genomic instability, also referred to as a mutator phenotype, a typical feature of HNPCC (Breivik and Gaudernack, 1999; Toft and Arends, 1998). Although the mutator defect that arises from an impaired MMR system can affect any DNA sequence, microsatellite sequences are particularly sensitive to MMR abnormalities. It has indeed been hypothesised that random accumulation of mutations in genes with repetitive DNA tracts within their coding regions, as a result of a defect in the MMR pathway, is associated with tumour progression (Barnetson *et al.*, 2000). The association between microsatellite instability and cancer is not restricted to HNPCC, but is best documented for this condition (Peltomaki, 2001a; Peltomaki, 2001b).

Animal models of MMR deficiencies have been generated (Table 1.4) and they have corroborated the involvement of MMR mutations in cancer development (Heyer *et al.*, 1999).

**Table 1.4. Phenotypes of MMR-deficient mice.**

Gene	Tumour phenotype	Meiotic phenotype	Microsatellite instability	Reference
<i>Mlh1</i>	Lymphomas GI <sup>a</sup> tumours	Male and female sterility	Tumours and normal tissues from homozygous null mice	Edelmann <i>et al.</i> , 1996; Prolla <i>et al.</i> , 1998
<i>Msh2</i>	Lymphomas GI and skin tumours	None	Tumours from homozygous null mice	Reitmair <i>et al.</i> , 1995
<i>Msh3</i>	No tumours until late age Lymphomas GI tumours	None	Tumours from homozygous null mice	de Wind <i>et al.</i> , 1999; Edelmann <i>et al.</i> , 2000
<i>Msh6</i>	GI tumours <sup>b</sup> None <sup>b</sup>	None	Not detected	de Wind <i>et al.</i> , 1999; Edelmann <i>et al.</i> , 1997
<i>Pms1</i>	None	None	Not detected	Prolla <i>et al.</i> , 1998
<i>Pms2</i>	Lymphomas	Male sterility	Tumours and normal tissues from homozygous null mice	Baker <i>et al.</i> , 1995; Prolla <i>et al.</i> , 1998

<sup>a</sup> GI: gastrointestinal

<sup>b</sup> Depending on the genetic background

Unlike adult HNPCC individuals, mice heterozygous for MMR defects do not develop cancer at an early stage (up to two years of age) and inactivation of both alleles is required for cancer development. This divergence may be accounted for by the limited life span of mice, relative to humans. *Msh2*- (Reitmair *et al.*, 1995) and *Msh6*-deficient mice (Edelmann *et al.*, 1997) are viable and fertile, but show a strong cancer predisposition phenotype. *Mlh1* (Edelmann *et al.*, 1996; Prolla *et al.*, 1998) and *Pms2* (Baker *et al.*, 1995; Prolla *et al.*, 1998) mice are also viable in

the homozygous state and show a predisposition to cancer, but exhibit varying degrees of sterility (Table 1.4), indicating that mutations in *Mlh1* and *Pms2* cause defects in meiosis, and suggesting unique roles for these genes in gametogenesis (Baker *et al.*, 1995; Edelman *et al.*, 1996; Prolla *et al.*, 1998). Microsatellite instability was observed in mice nullizygous for *Msh2* (Reitmair *et al.*, 1995), *Msh3* (de Wind *et al.*, 1999; Edelman *et al.*, 2000), *Mlh1* (Edelman *et al.*, 1996; Prolla *et al.*, 1998) and *Pms2* (Baker *et al.*, 1995; Prolla *et al.*, 1998). In contrast, microsatellite instability has not been reported in *Msh6*- or *Pms1*-deficient animals (Baker *et al.*, 1995; Edelman *et al.*, 1996; Prolla *et al.*, 1998).

### 1.5.3.2. MMR proteins in cell cycle surveillance and apoptosis

MMR proteins appear to be involved in biological activities, other than simple mismatch repair, including signalling of apoptosis or cell cycle arrest in response to DNA damage (Hickman and Samson, 1999; Toft *et al.*, 1999; Wu *et al.*, 1999). The futile-repair model proposes that repetitive, yet unsuccessful, cycles of DNA mismatch repair events directed to the newly synthesized strand opposite a major DNA lesion (*e.g.* induced by oxidative stress or radiation), create persistent DNA strand breaks and gaps without removing the offending DNA adducts. Futile cycles of excision and resynthesis result in the generation of persisting DNA termini, which might trigger an apoptotic signal that leads to cell death. Alternatively, cell death may be due to the binding of MMR proteins to DNA adducts, which either leads directly to a checkpoint response, or blocks other DNA transactions such as replication, transcription and proper damage repair (Aquilina and Bignami, 2001). MMR mutations may therefore predispose to malignancy not only through failure to repair mismatched DNA lesions, but also through failure to engage apoptosis and eliminate damaged cells that otherwise may be potential founders of clones bearing unrepaired mutations. The obvious implication of this finding is the refractory behaviour of MMR-deficient tumours to chemotherapy (Toft *et al.*, 1999).

## 1.6. Mechanistic models of trinucleotide repeat expansion

The sequence and location of the expanded trinucleotide repeat are most likely responsible for the unique mechanism of pathogenesis characteristic of each disorder. Nevertheless, it is not unreasonable to speculate that a common, or a least similar, mutation mechanism could be shared by different trinucleotide repeat sequences. It is currently not clear exactly how the triplet repeat expansion occurs, why the mutation frequency depends on repeat length and why some diseases display a sex bias. It is however generally assumed that unusual structural features of the repeats, play a role, and several models for expansion have been proposed, involving alternative DNA structures in erroneous DNA replication, recombination or repair.

Trinucleotide repeat sequences possess sequence motif and symmetry elements that allow increased flexibility of the double DNA helix and the formation of stable and reproducible

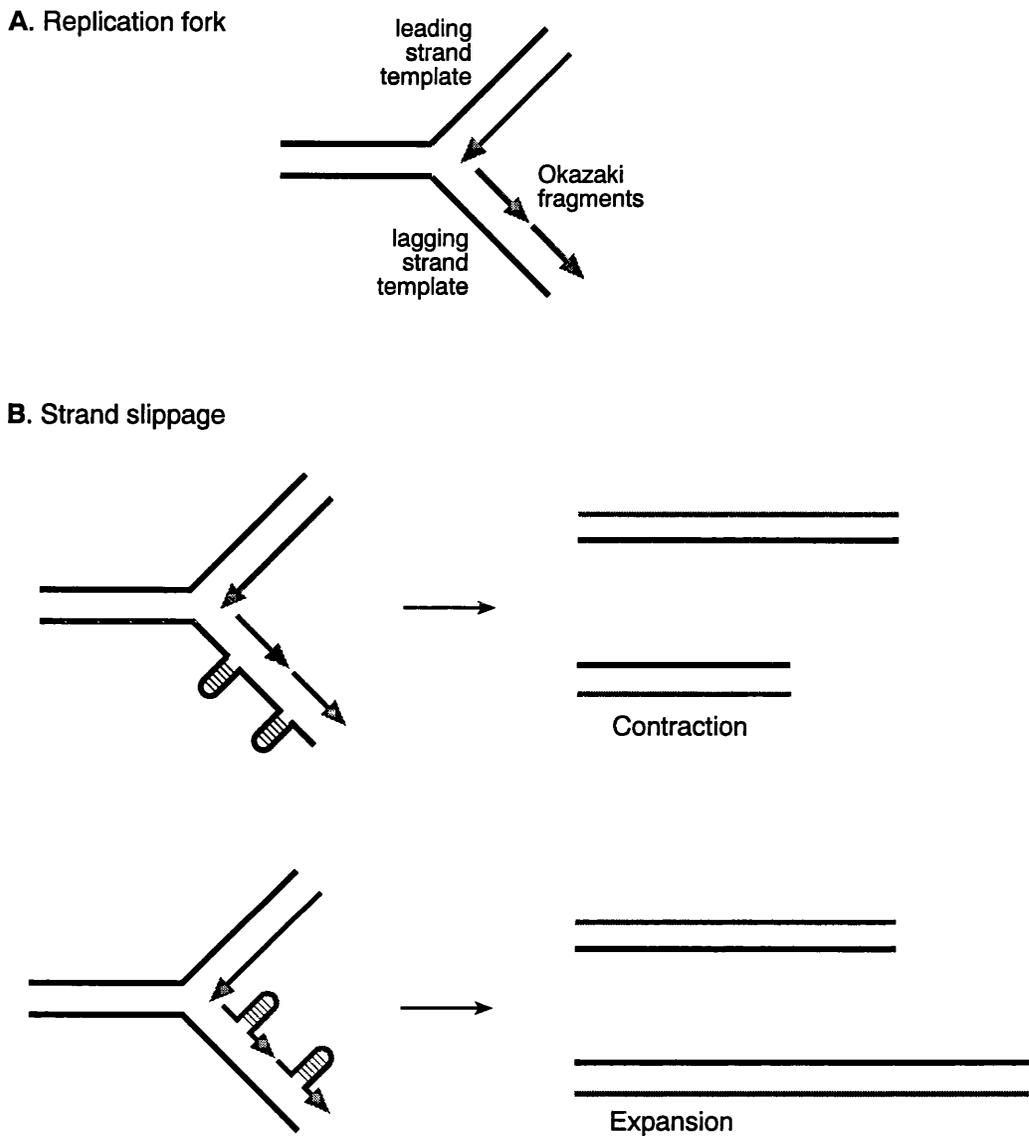
structures, alternative to the more conventional double-stranded B-DNA conformation. Indeed, several studies of single-stranded and cloned trinucleotide repeats suggest the formation of hairpins (Darlow and Leach, 1998a; Darlow and Leach, 1998b; Petruska *et al.*, 1996; Zheng *et al.*, 1996), triplexes (Bidichandani *et al.*, 1998; Gacy *et al.*, 1998; Sakamoto *et al.*, 1999), slipped-stranded structures (Pearson *et al.*, 1998a; Pearson *et al.*, 1997; Pearson and Sinden, 1996; Pearson and Sinden, 1998a) and parallel duplexes (LeProust *et al.*, 2000). The unique structural features of trinucleotide repeat sequences might therefore affect their genetic stability, and their presentation to other interacting molecules, particularly proteins involved in DNA metabolism (the structural properties of trinucleotide repeat sequences are discussed in further detail in Chapter 4).

Current models for expansion fall into two basic classes: those that are replication based, and those that involve recombination. It is important to be aware that most of these mechanisms have been proposed based on studies performed in simple model organisms, such as *E. coli* and *S. cerevisiae*. Unlike humans, microbial systems have not mimicked the inclination towards repeat expansion (Freudenreich *et al.*, 1997; Kang *et al.*, 1995b; Schweitzer and Livingston, 1997; Wells *et al.*, 1998). This disparity could reflect fundamental differences in DNA metabolism, and it raises serious questions about the extent in which the findings based on simple model organisms apply to humans.

### **1.6.1. DNA polymerase slippage model**

In simple strand slippage models, transient dissociation of the nascent strand from the template during DNA replication provides an opportunity for the strand to slip relative to one another, and for primer template misalignment during replication of a triplet repeat tract. If the nascent strand slips backwards and DNA polymerase primes DNA synthesis from this position, a gain of repeat units or expansion may result, if the loop-out is not subsequently repaired by the MMR system prior to a second round of DNA replication. If, on the other hand, the nascent strand slips forward, a deletion can result (Richards and Sutherland, 1994; Wells, 1996) (Figure 1.4).

The tendency for strand misalignment during replication of triplet repeat regions is attributed to the ability of primer or template strand to fold into hairpin structures stabilised by intrastrand base-pairs. Such stable hairpins may act as DNA polymerisation obstacles, impeding the transversal of the replication fork through a repeat stretch and creating the opportunity for strand misalignment. *In vitro* DNA polymerisation is indeed blocked through trinucleotide repeat tracts (Kang *et al.*, 1995c; Ohshima and Wells, 1997). *In vivo*, CGG•CCG and GAG•CTG tracts block the progression of the replication fork in bacteria (Samadashwily *et al.*, 1997), while GAA•TTC tracts inhibit replication in mammalian cells (Ohshima *et al.*, 1998). In addition, DNA replication of methylated CGG•CCG expansions, in lymphoblastoid cell lines derived from FRAXA patients, occurs late in S phase (Hansen *et al.*, 1993). Interestingly, interruptions within the repetitive DNA sequences, which destabilise possible alternative structures, abolish replication blockage (Samadashwily *et al.*, 1997). These findings suggest that the formation of unusual DNA structures



**Figure 1.4. DNA polymerase slippage model for trinucleotide repeat expansion.**

The parental DNA sequence is shown in black. Arrows indicate the DNA strands being synthesised, with the arrowheads pointing in the direction of DNA polymerase movement. **(A)** Schematic representation of the replication fork. While the leading strand is synthesised in a continuous manner, replication of the lagging strand depends on the discontinuous synthesis of Okazaki fragments. **(B)** Although replication slippage can in theory occur during both leading and lagging strand synthesis, the latter is probably more prone to polymerase slippage. Strand slippage is facilitated by the formation of alternative DNA structures, presumably stabilised by intra-strand base pairs. Slippage within the template DNA strand will lead to repeat contraction, whereas slippage within the newly synthesised DNA strand will lead to repeat expansion, if alternative secondary DNA structures are not properly repaired prior to the following genome replication event.

accounts for the observed replication blockage, which might facilitate misalignment between the newly synthesised and the template DNA strand, potentially leading to changes in repeat number.

Expansions in bacteria and yeast occur predominantly when the strand that forms more stable secondary structure is the lagging daughter strand, while deletions occur predominantly when this strand is the template for lagging strand DNA synthesis (Freudenreich *et al.*, 1997; Kang *et al.*, 1995b; Maurer *et al.*, 1996; Miret *et al.*, 1997). Discontinuous synthesis of the lagging strand implies that a portion of the lagging strand must be transiently single-stranded, so that a trinucleotide repeat has better chance to form a secondary structure in the lagging strand template. The greater stability of the hairpins formed by the single-stranded CTG strands compared to the CAG strands (Mitas, 1997) is sufficient to explain the observed differences in deletion frequencies between leading and lagging strand (Kang *et al.*, 1995b). The instability of GAA•TTC repeats is also greater when GAA is the lagging strand template, than when it is the leading strand template (Ohshima *et al.*, 1998). This differential instability may again be explained by the greater ability of single-stranded GAA to adopt a more stable secondary DNA structure than single-stranded TTC sequences (LeProust *et al.*, 2000). Cloning TGG•CCA into *E. coli* yields similar results: higher levels of trinucleotide repeat instability are detected when the TGG sequence serves as template for the lagging strand synthesis, which is consistent with the higher G content of this repeat, enabling it to form quadruplex structures (Pan and Leach, 2000). The proposed mechanism is therefore based on the discontinuous synthesis of the lagging strand during DNA replication (Figure 1.4).

The replication slippage model could account for relatively small changes in repeat number, and could also explain how interspersed interruptions stabilise the repeat tract. For large expansions to occur via this mechanism, many slippages would be needed in a single round of replication. This might be possible in *E. coli*, if the Okazaki fragments generated during lagging strand DNA synthesis were entirely comprised within the repeat, so that they are free to slip repeatedly during replication. Such fragments are not anchored by neighbouring non-repetitive sequences and might be able to slip or slide during DNA polymerisation, possibly resulting in addition of many copies of the repeat to the lagging strand (Sarkar *et al.*, 1998).

The replication-based model, in which slippage of perfect repeat Okazaki fragments leads to repeat expansion, has been favoured in the literature, however DNA polymerase slippage has not been definitively proven to be responsible for the trinucleotide expansions found in human disorders.

#### 1.6.1.1. Transcription interference in repeat length changes

Since in all situations in which the expansion of a trinucleotide repeat is associated with a human disease, the repetitive sequence is located within a gene, there is the possibility for transcription-associated events to influence the stability of trinucleotide repeats, by interfering with the transversal of the replication fork through a repetitive sequence. Active transcription through triplet repeat tracts increases the frequency of deletions of long CAG•CTG tracts from plasmids in *E. coli*, and reduces the frequency of expansions (Bowater *et al.*, 1997; Bowater *et al.*, 1996;

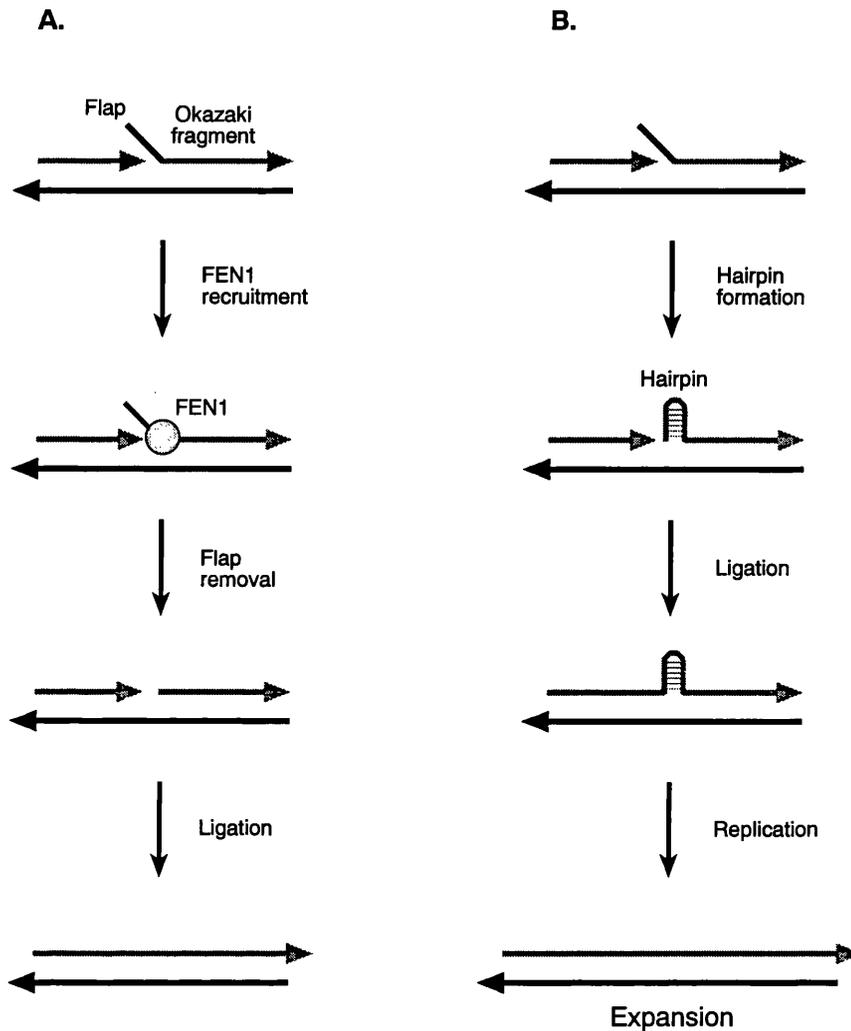
Schumacher *et al.*, 2001). In contrast, transcription fails to affect the stability of shorter CTG•CAG repeat tracts (Schmidt *et al.*, 2000). It has been hypothesised that transcription of long CTG•CAG repeats influences the transition between stationary and exponential growth phase, conferring a growth advantage to cells harbouring deleted CTG•CAG tracts (Bowater *et al.*, 1997). The selective growth advantage, if any, would not be so evident in bacteria carrying shorter trinucleotide arrays, therefore reconciling these two apparently contradictory observations.

The presence of transcription complexes within trinucleotide repeat sequences may favour the formation of secondary structures, enhancing the error rate of replication, leading to a higher frequency of deletions (Bowater *et al.*, 1997). Alternatively, DNA replication and transcription can occur simultaneously on the same DNA sequence, leading to eventual head-on collisions of the RNA polymerase with the replication machinery. Head-on collisions could lead to an increase in double strand breaks and illegitimate recombination (Schumacher *et al.*, 2001).

By analogy with factors that affect repeat instability in bacteria, transcription through a given region may mediate the mutation mechanism in mammals. The tissue-specific pattern of somatic mosaicism detected in SBMA patients appears to correlate with the expression pattern of the androgen receptor, as assessed by western blot and immunocytochemistry analysis (Tanaka *et al.*, 1999), suggesting the involvement of gene expression and transcription in the mechanism of expansion. An effect of transgene expression was suggested in transgenic mice for *HD* exon 1, given that the lowest levels of somatic mosaicism were detected in a line that did not express the transgene (Mangiarini *et al.*, 1997). Nevertheless, tissue-specific levels of somatic instability in transgenic mice carrying repeat tracts, derived from the human *DM1* locus, did not show a simple correlation with the transcriptional levels of the transgene. Although the transgenic repeat was indeed ubiquitously expressed in the transgenic lines exhibiting somatic mosaicism, the highest expression levels, as assessed by reverse transcriptase polymerase chain reaction (RT-PCR) analysis, did not correlate with the tissues showing the greatest somatic instability (Fortune, 2001; Lia *et al.*, 1998).

#### 1.6.1.2. FEN1 interference model

Attractive as it might be, the replication slippage model cannot easily explain the clear expansion-biased nature of trinucleotide instability. Therefore, yet another model based on the displacement synthesis of Okazaki fragments has been proposed (Gordenin *et al.*, 1997). Each replication fork comprises a continuous leading strand and a discontinuous lagging strand, which consists of several fragments. Unlike continuous synthesis at the leading strand, which is maintained mostly as a duplex, the lagging strand synthesis proceeds in a discontinuous fashion. A portion of the lagging strand template must be rendered single-stranded to permit the priming of an Okazaki fragment. As a consequence of the discontinuous mechanism of DNA synthesis of the lagging strand, an Okazaki fragment may have its 5' end displaced by polymerase extension of the immediately upstream Okazaki fragment. Normally, the processing of Okazaki fragments involves the excision of the flap by flap endonuclease 1 (FEN1), gap filling and ligation (Figure 1.5).



**Figure 1.5. A model of repeat expansion via a FEN1-resistant folded flap.**

(A) Usually, the displaced 5' flap of an Okazaki fragment is removed by the structure-specific endonuclease FEN1. Synthesis of the upstream Okazaki fragment results in the displacement of the 5' end of the downstream Okazaki fragment to generate a flap. FEN1 is loaded at the 5' end of the flap and slides down to the base of the flap, removing the flap from the Okazaki fragment. (B) Depending on its sequence, a single-stranded flap may form FEN1-resistant structures, such as hairpins. Ligation and replication results in expansion, by insertion of the FEN1-resistant flap into the new double-stranded DNA helix.

However, a displaced flap containing CAG or CTG repeats, or even other triplet repeats, may form a thermodynamically favourable secondary structure. Given that FEN1 can only act on single-stranded DNA, a partially double-stranded flap blocks FEN1 activity, and hence its own removal (Chen *et al.*, 1995; Gacy *et al.*, 1995; Spiro *et al.*, 1999). Aberrant flaps, resulting from the formation of a hairpin or triplex structure, will be inefficiently excised on completion of lagging-strand replication. Ligation of the 5' end of a flap (which has not been removed by FEN1) to the 3' end of the upstream Okazaki fragment will result in a nascent DNA strand with an increased number of repeats equal to the length of the flap, and eventually result in mutation, with lack of excision providing the initial step in generating large DNA expansions (Gordenin *et al.*, 1997). It was demonstrated that in yeast, secondary structure could indeed inhibit flap processing in a length-dependent manner (Henricksen *et al.*, 2000; Spiro *et al.*, 1999). In addition deletion of *rad27* (the yeast homologue of human *FEN1*), causes length-dependent destabilisation of the CAG•CTG tracts and a substantial increase in expansion frequency (Freudenreich *et al.*, 1998). The FEN1 interference model is appealing because it may account for the gross mutation bias towards expansion.

### **1.6.2. Recombination-dependent mechanisms of expansion**

Theoretically, trinucleotide repeat expansion could also result from recombination-dependent events. Complex patterns of haplotype variation and gene disequilibrium in FRAXA were explained on the basis of unequal crossing-over or gene conversion (Brown *et al.*, 1996; Losekoot *et al.*, 1997). In DM1, there is no direct evidence for exchange of flanking material around the CTG•CAG repeat (Imbert *et al.*, 1993; Jansen *et al.*, 1992a; Shutler *et al.*, 1994), with only one exception being reported (O'Hoy *et al.*, 1993). In addition, GAA•TTC instability in FRDA is not a function of the number of long alleles, ruling out homologous recombination as a major mechanism of mutation (Gacy *et al.*, 1998). The usual mechanisms for homologous recombination appear not to be essential for triplet repeat mutation also in *E. coli*, since expansions have been observed in *recA*<sup>-</sup> strains (Kang *et al.*, 1995b; Kang *et al.*, 1996; Napierala *et al.*, 2002; Ohshima *et al.*, 1996a; Ohshima *et al.*, 1996b) and similar extents of deletions were observed in isogenic strains with functional or deficient RecA activity (Jaworski *et al.*, 1995).

However, the expansion of CTG•CAG trinucleotide repeats by recombination, without the exchange of flanking markers, has been reported in recombination-proficient *E. coli* strains, transformed with two distinct plasmids carrying repetitive sequences (Jakupciak and Wells, 1999; Jakupciak and Wells, 2000), as well as in *S. cerevisiae* (Jankowski *et al.*, 2000; Richard *et al.*, 1999). Moreover, conditions that favour recombination events increase the frequency of expansion mutations (Jakupciak and Wells, 1999; Jankowski *et al.*, 2000; Richard *et al.*, 2000), in contrast to the massive deletion-biased instability previously detected in microbial systems (Freudenreich *et al.*, 1997; Kang *et al.*, 1995b; Schweitzer and Livingston, 1997; Wells *et al.*, 1998). This finding led to the hypothesis that, under especial circumstances, complex recombination events may be a robust process to originate trinucleotide repeat length variability in bacteria and in yeast. It has

therefore been speculated that similar complex recombination events, between sister chromatids or homologous chromosomes, could also play a role in both somatic and germline trinucleotide repeat instability in humans (Jakupciak and Wells, 1999; Jakupciak and Wells, 2000), similar to the non-reciprocal transfer of repeats between alleles in the germline, previously described for human minisatellites (Jeffreys *et al.*, 1994).

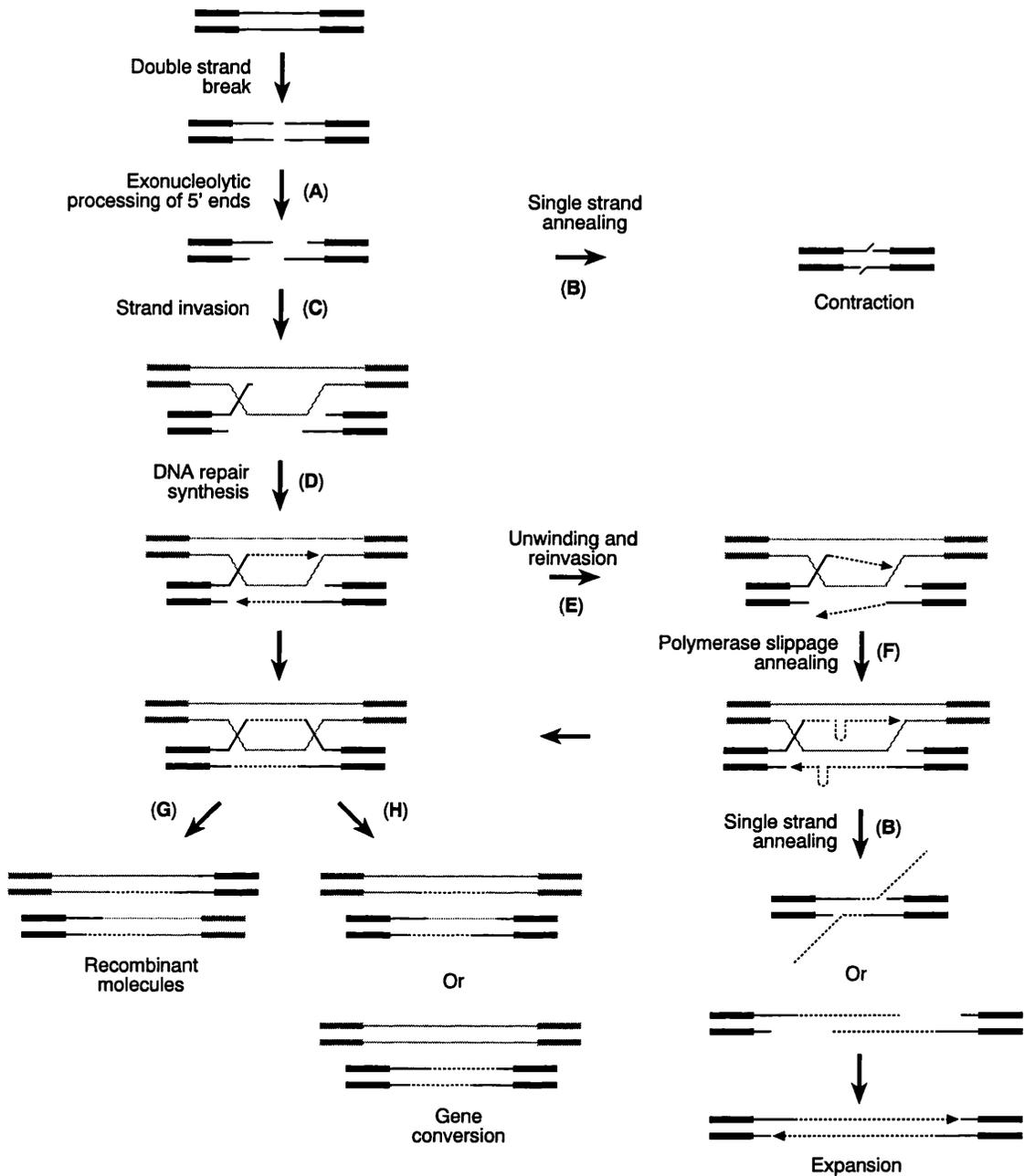
#### **1.6.2.1. Generation of DNA strand breaks within trinucleotide repeat sequences**

Recombination-mediated trinucleotide repeat instability requires DNA strand breaks. Replication stalling at CAG•CTG tracts (Kang *et al.*, 1995c; Ohshima and Wells, 1997; Samadashwily *et al.*, 1997), may increase the rate of double strand break formation within repeat sequences, and stimulate recombination events (Michel, 2000). Indeed, CAG•CTG repeats showed length-dependent susceptibility to strand breaks in yeast (Freudenreich *et al.*, 1998). Alternatively, the presence of unprocessed Okazaki fragments could also generate recombinogenic double strand breaks (Gordenin *et al.*, 1997).

#### **1.6.2.2. Trinucleotide repeat instability by double strand break repair**

According to this model, trinucleotide repeat instability would be initiated by an intramolecular double strand break. Exonucleolytic cleavage of the 5'-containing strands leaves two 3' over-hang ends, which can be processed in a number of different ways, leading to changes in the repeat number. Annealing of the two 3'-containing complementary strands, followed by DNA ligation would result in repeat contraction (Figure 1.6). Alternatively the gaps formed provide binding sites for the recombination machinery. Both 3' ends would each invade the donor template sequence, on a sister chromatid or sister chromosome, which might be aligned out-of-register with the first chromatid or chromosome. DNA repair synthesis extends both chains, and both the recipient and the donor locus become heteroduplexes, with one original and one newly synthesised strand (Figure 1.6). By analogy to the normal replication of trinucleotide repeat sequences, it might be imagined that during repair synthesis DNA polymerase undergoes strand slippage, facilitated by the formation of hairpin structures, to further add or subtract repeats (Paques *et al.*, 1998). In fact, DNA synthesis associated with double strand breaks is more prone to errors than S phase genomic DNA synthesis (Richard *et al.*, 1999). Resolution of the junctions may or may not imply exchange of non-repetitive flanking sequences, resulting in the generation of two recombinant DNA molecules or in a gene conversion-type mechanism (Figure 1.6). The exact repeat length gained will depend on the degree of horizontal slippage during the alignment step, between the 3' free end and the template DNA strand (Jankowski *et al.*, 2000).

Alternatively the newly synthesised DNA does not remain base-paired to the template, as in normal DNA replication, instead, newly synthesised strands are displaced and reannealed, in a process called synthesis-dependent strand annealing (SDSA) (Paques *et al.*, 1998; Richard and Paques, 2000) (Figure 1.6). DNA synthesis is thus conservative and fundamentally different from the semiconservative genome replication that occurs during the S phase of the cell cycle. The



**Figure 1.6. Expansion of trinucleotide repeats by double strand break repair.**

A double strand break occurs within the trinucleotide repeat (thin lines) of the recipient molecule (black). Exonucleolytic cleavage of the double strand break creates 3' ends (A). Single strand annealing of tandem repeats, followed by DNA ligation, leads to repeat contraction or expansion (B). Strand invasion of the homologous template (grey) by a resected 3' end initiates recombination and forms two Holliday-like junctions (C). DNA repair synthesis occurs on both strands (dashed lines) and restores the double-stranded trinucleotide sequence (D). Arrowheads indicate the direction of DNA synthesis. Unwinding and out-of-register reinvasion of one (or both) the newly synthesised strand(s) may happen during repair synthesis (E), as well as strand misalignment, involving single-stranded loops (F). Different ways of resolving the junctions generate distinct products. Exchange of flanking non-repetitive sequences (thick lines) results in recombinant molecules (G). Alternatively gene conversion or synthesis-dependent strand annealing type mechanisms leave the flanking regions unaltered (H)

recipient locus will receive two newly synthesised strands of DNA, while the donor template remains unchanged (Paques *et al.*, 1998). Since alignment of complementary newly synthesised strands can occur in different registers, the total number of repeats in the recipient site can easily have fewer or more copies than the donor template. In conjunction, double strand break repair and replication slippage, could indeed mediate multiple-fold expansions of trinucleotide repeat sequences (Richard and Paques, 2000).

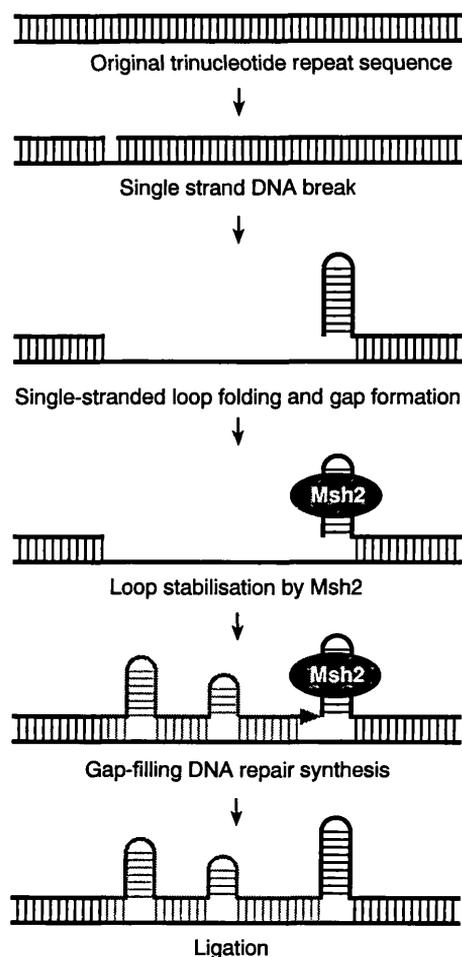
In summary, it is theoretically possible, and it has indeed been experimentally confirmed, that under controlled circumstances, trinucleotide repeat instability in simple organisms, can arise through a variety of molecular pathways, such as replication slippage, transcription-dependent mechanisms, and complex recombination events, initiated by double strand breaks. However, it is not clear if these mechanisms are operating in mammalian cells, and if so, to which extent they are responsible for trinucleotide repeat mutation.

### **1.6.3. Models of germline trinucleotide repeat instability**

Following the repeat dynamics in the male germline of transgenic R6/1 mice, carrying an expanded *HD* exon 1 (Mangiarini *et al.*, 1996), trinucleotide repeat expansion in haploid cells does not appear to occur during either mitotic or meiotic replication. Rather, repeat instability appears to be a post-meiotic event, which occurs late in spermatogenesis, when haploid spermatids undergo terminal differentiation into mature spermatozoa, in the absence of DNA replication or recombination (Kovtun and McMurray, 2001). The parallel between the age of the transmitting mother and the enhanced deletion-biased germline instability, described in other mouse models (Kaytor *et al.*, 1997; Sato *et al.*, 1999), is also consistent with a mutation process that occurs in a developmental time window of the oocyte, following meiotic replication, but prior to fertilisation (Kaytor *et al.*, 1997).

The model proposed by Kovtun and McMurray hypothesises that once DNA strand breaks form, single-stranded folding generates loops, possibly stabilised by Msh2 protein, leaving a gap behind. The gap is subsequently filled by repair-associated DNA synthesis, which may also create additional loop-outs and hairpins by polymerase slippage within the repeat tract (Figure 1.7). Following ligation, the loops would be trapped in the DNA (Kovtun and McMurray, 2001). If not repaired by the DNA repair machinery, heteroduplex sequences, similar to the slipped-stranded intermediate DNA (SI-DNA) structures previously described (Pearson *et al.*, 1997), would be present in sperm, possibly giving rise to mosaic embryos following fertilisation.

In support of this model *Msh2*-deficiency completely abolishes germline expansion, suggesting that *Msh2* may indeed mediate germline instability (Kovtun and McMurray, 2001). It was further suggested that a similar model might also account for the somatic instability observed not only in R6/1 mice but also in humans, consistent with previous findings that *Msh2* deletion stabilises expanded CAG•CTG repeats in the soma of R6/1 mice (Manley *et al.*, 1999b).



**Figure 1.7. Mechanism for trinucleotide repeat expansion in germ cells by gap repair.**

Introduction of a single strand break enables single stranded folding into a secondary structure such as loop or hairpin, which is stabilised upon Msh2 recognition and binding. Gap-filling DNA repair synthesis fills the gap and creates the opportunity for polymerase slippage and formation of additional loops. Loops are trapped into the double-stranded DNA molecule following ligation (adapted from Kovtun and McMurray, 2001).

In contrast with the model proposed by Kovtun and McMurray, the high mutation frequency estimated for the male germline in HD patients, appears to be more consistent with a mutational process that occurs throughout germline mitotic division in the renewing spermatogonial stem cell population, rather than resulting from a single event (Leeflang et al., 1999). Recombination between the HD-containing sister chromatids is another theoretical possibility, but appears highly unlikely as it would have to occur in almost every meiosis and result in an increase in repeat number on both sister chromatids, in order to account for an average mutation frequency of 98% in individuals with at least 50 repeats (Leeflang et al., 1999). No mechanism consistent with these assumptions has yet been described. In addition, simple mathematical modelling of germline instability, always involving small repeat changes, in mice carrying a single copy of the 3'UTR of the DMPK gene (Monckton et al., 1997), taking into

account the number of cell divisions of spermatogenic stem cells and the calculated mutation frequency, provided support for a model of germline repeat instability based on replication slippage events and on the displacement of Okazaki fragments (Zhang et al., 2002). Nevertheless, time *per se*, rather than the number of cell divisions, could also serve as the critical factor controlling repeat expansions in spermatogenic stem cells in a cell division-independent mechanism (Zhang et al., 2002).

## 1.7. Project design

Most trinucleotide repeat diseases are associated with the expansion of CAG•CTG trinucleotides, including DM1, HD, SBMA, DRPLA and SCAs types 1, 2, 3, 7 and 8 and 17 (Section 1.1). Although progress has been made in understanding the pathophysiology of diseases associated with genetic instability of simple trinucleotide repetitive DNA sequences, the mechanisms of expansion of repeat sequences remain unclear. Trinucleotide repeat disorders may develop through a complex interplay between hereditary factors and the genomic instability induced by DNA damage within the cell environment. Much of this damage results from the intrinsic chemistry of DNA in living cells, and it is conceivable that the cellular processes designed to maintain genomic integrity will affect trinucleotide dynamics. Many questions about the mechanism and the consequences of expansion remain unanswered. These questions are not simply of academic interest, since their answers may suggest new therapeutic approaches to this group of diseases.

The detailed analysis of the dynamics of trinucleotide repeat mutation might be envisaged as a preliminary step in determining the common molecular mechanisms shared by trinucleotide repeat disorders. The transgenic *Dmt-D* line replicates the gross age-dependent, tissue-specific, expansion-biased somatic mosaicism observed in DM1 patients (Fortune *et al.*, 2000). In terms of the rate of repeat length mutation change as a function of time, it may even be greater in *Dmt-D* mice than in humans. These animals provide a valuable source of tissue samples, from which cell cultures might be established. Cell lines containing expanded CAG•CTG repeats may offer a powerful system, in which repeat dynamics can be accurately monitored under controlled conditions over large numbers of cell divisions in a relatively short period of time. The repeat instability mechanisms (*e.g.* mitotic vs. repair synthesis) could be easily tested in cell lines, which accumulate repeat length mutations over time. The establishment of a *Dmt-D* cell culture system and detailed characterisation of repeat dynamics *in vitro* should therefore shed more light on the molecular mechanisms involved in repeat length variation. In addition, *Dmt-D* cell lines may even create new avenues for the preliminary assessment of therapeutical intervention, by means of modulating trinucleotide repeat instability.

A major issue is the understanding of the pathway that links genotype to phenotype, and to distinguish between cause and consequence. Although each triplet repeat disorder contains a mutation in a single gene, the mechanism for each disorder is complex and requires identification

of all components of the triplet repeat pathway. Trinucleotide repeat expansions associated with human diseases, appear to lead to pathology by a variety of different mechanisms, including induction of methylation and gene silencing, inhibition of transcription elongation by higher-order DNA non-orthodox structures, aberrant mRNA processing, or a toxic gain-of-function at the protein level (Section 1.2). Consequently, therapeutical strategies aimed at preventing the pathophysiological consequences of trinucleotide expansions will most likely be highly specific, and only suitable to a singular disease or to a particular subclass of triplet repeat disorders. Regulation of gene expression, would appear as a possible therapeutical approach. In mice transgenic for tetracyclin-regulated *HD* exon 1, neurological signs are largely reversed when transgene expression is then turned off (Yamamoto *et al.*, 2000), suggesting that some neurological damage can indeed be reversed, by down-regulating the expression of the affected gene. However, conditional inactivation of *Hdh* expression in the forebrain of mice, at postnatal day five, led to neurological deficits, progressive neurodegeneration, motor phenotypes, and early mortality, suggesting that huntingtin is required for neuronal function and survival in the brain, and that a loss of function may partially contribute to HD pathogenesis (Dragatsis *et al.*, 2000). A similar scenario may also apply to other dominant polyglutamine disorders, rendering therapeutical strategies based on gene down-regulation less than ideal, if not specifically targeted to the expanded allele.

Alternatively, the development of means of causing repeat length contraction and reversion to non-pathogenic repeat sizes would be predicted to be potentially curative, possibly leading to reversion of symptoms, or at least interruption of disease progression. Nevertheless, suppressing trinucleotide repeat instability, resulting in decreased expansion rates, may also be highly beneficial, as it may possibly decelerate clinical progression or delay disease onset. *Dmt-D* cell lines, which recreate the complex metabolism of trinucleotide repeats, offer the unique opportunity to assess the effect of multiple genotoxic agents on the dynamics of expanded CAG•CTG tracts. Potential chemical reagents, which might modify the somatic repeat dynamics, could be tested in cell culture models. In the long run, the preliminary identification of genotoxic modifiers of triplet repeat instability may suggest novel therapies based upon control of repeat size variation *in vivo*.

In summary, this project was designed, not only to achieve a better understanding of the molecular mechanisms underlying trinucleotide repeat mutation, but also to gather preliminary data supporting intervention at the DNA level as a valid therapeutical approach to battle trinucleotide repeat disorders.

## 2. Materials and methods

### 2.1. General materials

#### 2.1.1. Chemicals and reagents

AnalaR and molecular biology grade chemicals were obtained from Fisher Scientific, Merck Ltd. (BDH Laboratory Supplies), Riedel-de Haën or Sigma Chemical Company Ltd. Other reagents obtained from alternative sources are mentioned in Table 2.1.

**Table 2.1 Chemicals and reagents**

<b>Chemicals and reagents</b>	<b>Supplier</b>
[ $\alpha$ - <sup>32</sup> P]dCTP (3000 Ci/nmol)	Amersham Pharmacia Biotech
30% (w/v) acrylamide/bis 29:1 (3.3% C)	BioRad
Acetic acid	Fisher Scientific U.K. Ltd.
Agar	Difco Laboratories
Agarose	Boehringer Mannheim
Apicidin	Calbiochem
Bacto <sup>®</sup> tryptone	Difco Laboratories
Bacto <sup>®</sup> yeast extract (without amino acids)	Difco Laboratories
Boric acid	Fisher Scientific U.K. Ltd.
Chloroform	Fisher Scientific U.K. Ltd.
Ethanol, absolute 100%, analytical reagent	Bamford Laboratories
ExpressHyb hybridization solution	Clontech
Formamide	Fluka
Hydrochloric acid	Fisher Scientific U.K. Ltd.
IPTG	Boehringer Mannheim
PMSF	Boehringer Mannheim
Sephadex <sup>®</sup> G-50	Amersham Pharmacia Biotech
Sodium dodecyl sulphate	Anachem
X-gal	Boehringer Mannheim

#### 2.1.2. Immunochemicals

Immunochemicals were obtained from Sigma Chemical Company Ltd.

### 2.1.3. General disposable plasticware

General disposable plasticware materials used during the course of this project are listed in Table 2.2 below.

**Table 2.2. Plastic materials and suppliers.**

<b>Plastic Materials</b>	<b>Supplier</b>
0.2 ml micro-tubes	ABgene®
15 ml centrifuge tubes, gamma irradiated	Sterilin
20 $\mu$ l, 200 $\mu$ l and 1 ml filter pipette tips	Rainin Instrument Co. Inc. and Greiner Labortechnik
200 $\mu$ l and 1 ml pipette tips	Sarstedt
50 ml centrifuge tubes	Sterilin
90 mm Petri dishes	Philip Harris Scientific
Thin-wall polycarbonate 96-well plates and thermosealers	Costar

### 2.1.4. Tissue culture disposable materials

Tissue culture disposable plastic materials are listed in Table 2.3 below.

**Table 2.3. Tissue culture materials and suppliers.**

<b>Disposable Materials</b>	<b>Supplier</b>
100 x 20 mm polystyrene tissue culture dishes	Corning
25 cm <sup>2</sup> tissue culture flasks	Costar and Iwaki
5, 10 and 25 ml pipettes	Corning
6-, 12- and 96-well tissue culture microplates	Iwaki
60 x 15 mm polystyrene tissue culture dishes	Corning
75 and 150 cm <sup>2</sup> tissue culture flasks	Iwaki
Cell scrapers	Sigma
Cryo 1°C freezing container	Nalgene
Immunocytochemistry 8-well chamber glass slides	Nalge Nunc

### 2.1.5. Enzymes

The enzymes used for DNA analysis are listed in Table 2.4, together with their suppliers, from where their reaction buffers were also obtained.

**Table 2.4. Enzymes and suppliers.**

<b>Enzyme</b>	<b>Supplier</b>
<i>Taq</i> DNA polymerase	Bioline
T4 DNA ligase	GibcoBRL Life Technologies
<i>Eco</i> RI	GibcoBRL Life Technologies
<i>Hind</i> III	GibcoBRL Life Technologies
Proteinase K	Boehringer Mannheim
RecA	New England Biolabs

### 2.1.6. Nucleic acids size markers and mass ladders

Following gel electrophoresis, DNA and RNA samples were sized and/or quantified using the following size markers or mass ladders listed in Table 2.5.

**Table 2.5. Nucleic acid size markers and/or mass ladders and their suppliers.**

<b>Size marker and/or mass ladder</b>	<b>Supplier</b>
1 kb <sup>+</sup> ruler	GibcoBRL Life Technologies
2.5 kb molecular ruler	BioRad
Amplisize ruler	BioRad
<i>Hae</i> III digested $\phi$ X174 DNA	New England Biolabs
High MW mass ladder	GibcoBRL Life Technologies
<i>Hind</i> III digested $\lambda$ DNA	New England Biolabs
Low DNA mass ladder	GibcoBRL Life Technologies
RNA ladder	GibcoBRL Life Technologies

### 2.1.7. Kits

The kits used during the course of this project are listed in Table 2.6.

**Table 2.6. Kits and suppliers.**

<b>Kit</b>	<b>Supplier</b>
5'-Bromo-3-deoxyuridine detection and labelling kit I	Roche Boehringer
Nucleon <sup>®</sup> BACC kit	Nucleon
PGEM <sup>®</sup> -Teasy vector system	Promega
QIAprep <sup>®</sup> miniprep kit	Qiagen
QIAquick <sup>®</sup> spin kit	Qiagen
Ready-to-go DNA labelling beads (-dCTP)	Amersham Pharmacia Biotech

### 2.1.8. Membranes and paper

Hybond<sup>™</sup>-N<sup>+</sup> and Hybond<sup>™</sup>-N nylon membranes were obtained from Amersham Pharmacia Biotech. Magna nylon transfer membrane was purchased from Osmonics. Biodyne b membrane was obtained from Flowgen.

Schleicher & Schuell supplied gel blotting paper. Saran wrap was obtained from Dow.

### 2.1.9. Photography and autoradiography

Agarose gels were photographed using a UVP gel documentation system 7500. Konica X-ray film and autoradiography cassettes were obtained from Genetic Research Instrumentation Ltd. Films were developed using an X-Ograph Compact X2 system (X-Ograph Ltd.). Both fixing and developer solutions were supplied by Kodak.

Cells were photographed using a Canon EOS300 camera and 400 ISO films.

### 2.1.10. Microscopes

A Nikon TNS phase contrast microscope and a Zeiss Axiovert S100 fluorescence microscope were used during the course of this project.

### **2.1.11. Spectrophotometer**

A Shimadzu UV-1201 UV-Vis spectrophotometer was used to estimate the concentration and purity of nucleic acids in aqueous solution.

### **2.1.12. DNA crosslinker**

A Stratagene<sup>®</sup> UV crosslinker 2400 was used during the course of this project, to fix DNA and RNA onto nylon membranes and also to irradiate cultured mouse cells with defined doses of ultraviolet (UV) light (wavelength 254 nm).

## **2.2. Experimental materials**

### **2.2.1. Animals**

The *Dmt-D* and *Dmt-E* transgenic mice used during the course of this project were generated by Darren G. Monckton (Monckton *et al.*, 1997). All *Dmt* mice were on an FVB/N genetic background, and kept at the Central Research Facility of the University of Glasgow.

C57/black 6 mice, transgenic for the disrupted murine *Pms2* gene (Baker *et al.*, 1995) were obtained from Michael Liskay (Department of Molecular and Medical Genetics, Oregon Health and Science University, Oregon, USA) and bred onto the FVB/N background.

### **2.2.2. Bacterial host strains**

*Escherichia coli* TOP10 strain was obtained from Invitrogen.

### **2.2.3. Vectors**

pGEM<sup>®</sup>-T Easy bacterial vector was obtained from Promega.

### **2.2.4. DNA sources**

pGEM-T750.19, pGEM-T750.21 and pGEM-T750.22 constructs, containing CTG•CAG trinucleotide repeats subcloned in pGEM<sup>®</sup>-T Easy bacterial vector, were a gift from Christine Howarth (Division of Molecular Genetics, University of Glasgow, Glasgow, UK). In all of them, a human CTG•CAG repeat derived from the *DMI* locus was amplified using oligonucleotide primers DM-H and DM-BR (Table 2.7) and the polymerase chain reaction (PCR) product was subsequently ligated into pGEM<sup>®</sup>-T Easy.

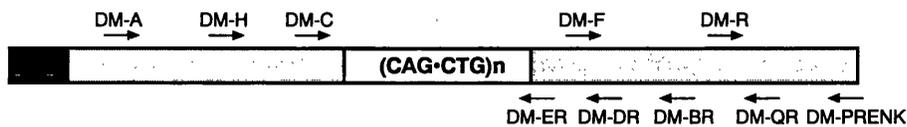
### 2.2.5. Oligonucleotides

The oligonucleotide primers used in this project were obtained from Genosys. The primer sequences are listed in Table 2.7 and the locations of the *Dmt* transgenic oligonucleotide primers are shown in Figure 2.1.

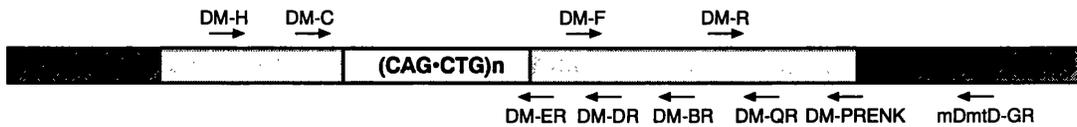
**Table 2.7. Oligonucleotides name, 5'-3' sequence, melting temperature ( $T_m$ ) and target sequence.**

Oligonucleotide	5'-3' Sequence	$T_m$	Target sequence
DM-A	AGTGCAGTTCACAACCGCTCCGAGC	67°C	CTG repeat 5' flanking region
DM-BR	CGTGGAGGATGGAACACGGAC	64°C	<i>Dmt</i> transgene
DM-C	AACGGGGCTCGAAGGGTCTCT	64°C	<i>Dmt</i> transgene
DM-DR	CAGGCCTGCAGTTTGCCCATC	64°C	<i>Dmt</i> transgene
DM-ER	AAATGGTCTGTGATCCCCC	60°C	<i>Dmt</i> transgene
DM-F	CTGAGGCCCTGACTGGGATGGGCAAACCTGC	72°C	<i>Dmt</i> transgene
DM-GR	GCAGGGCGTCATGCACAAGAAA	62°C	<i>Dmt</i> transgene
DM-H	TCTCCGCCAGCTCCAGTCC	66°C	<i>Dmt</i> transgene
DM-PRENK	GTCCGGTACCGAATTCGGCTAGCTCCTCCC AGACCTTC	73°C	<i>Dmt</i> transgene
DM-QR	CACTGTGGAGTCCAGAGCTTTG	62°C	<i>Dmt</i> transgene
DM-R	GTCCCTCCGACTCGCTGACAG	64°C	<i>Dmt</i> transgene
H1	TGCTTACCTTGTTACGACTTA	57°C	Mouse mitochondrial DNA
L1	CGCTCTACCTCACCATCTCTT	62°C	Mouse mitochondrial DNA
mDmtD-GR	AAAGGCAGGCATGGTTAGATTGAC	61°C	Transgene 3' flanking region on mouse genomic DNA
mP2-1	TTCGGTGACAGATTTGTAAATG	55°C	Mouse <i>Pms2</i> gene
mP2-2	TTTACGGAGCCCTGGC	57°C	Mouse <i>Pms2</i> disrupted gene
mP2-3	TCACCATAAAAAATAGTTTCCCG	55°C	Mouse <i>Pms2</i> gene
mUSF-A	GCCCCTGCCTCACCGTATAG	63°C	Mouse <i>upstream stimulatory factor 2</i> gene
mUSF-BR	CTGGGGTCCACCACTTCAAG	62°C	Mouse <i>upstream stimulatory factor 2</i> gene

### A. Human genomic DNA at the *DM1* locus



### B. *Dmt-D* transgene



- *DMPK* coding DNA
- Sequence of mouse genomic DNA on chromosome 11
- *Dmt162* transgene
- CAG-CTG repetitive tract

**Figure 2.1. Annealing sites for the oligonucleotide used to amplify the *Dmt162* transgene.** (A) The figure represents a fragment of the human *DM1* locus, from which the *Dmt162* transgene was derived. (B) In the *Dmt-D* mouse line, the transgene integrated on mouse chromosome 11. The figure also shows the annealing sites for the primers used during the course of this work.

## 2.2.6. Probes

Double-stranded DNA probes used in Southern blot hybridisations included (a) DM56, a CTG•CAG repeat PCR product amplified with oligonucleotide primers DM-C and DM-ER; and (b) DM-F/DM-PRENK, a non-repeat *Dmt* transgenic PCR product. Both were generated by PCR amplification and purified using the Qiagen PCR purification kit.

Double-stranded DNA probes, generated by PCR amplification of mouse genomic DNA samples and cloned into TOP10 plasmids, were used in northern blot hybridisation to detect (a) *cytochrome oxidase complex II* and (b)  $\beta$ -*actin* mRNA transcripts. Both probes were kindly provided by Chung-Mei Chen (Division of Molecular Genetics, University of Glasgow, Glasgow, UK).

## 2.3. Solutions

### 2.3.1. General solutions

$\lambda$  DNA x *Hind*III and  $\phi$ x174 DNA x *Hae*III size ladders

25 ng/ $\mu$ l  $\lambda$  DNA digested with *Hind*III and 25 ng/ $\mu$ l  $\phi$ x174 DNA digested with *Hae*III, 1X TE, 1X DNA loading dye.

**0.5M EDTA pH 8.0**

0.5M EDTA with NaOH to pH 8.0.

**1 kb+ ladder**

60 ng/μl 1 kb ladder, 1X DNA loading dye in 1X TBE.

**10% (w/v) SDS**

10 % (w/v) SDS in H<sub>2</sub>O.

**100X Denhardt's solution**

2% (w/v) Ficoll<sup>®</sup>400, 2% (w/v) polyvinylpyrrolidone, 2% (w/v) BSA.

**100X Ethidium bromide-acridine orange stock solution**

2.75 mM ethidium bromide, 1.25 mM acridine orange, 2% (v/v) ethanol in H<sub>2</sub>O.

**10X MOPS pH 7.0**

0.2 M MOPS, 50 mM sodium acetate, 10 mM EDTA with NaOH to pH 7.0 (stored at 4°C in the dark).

**10X MOPS pH 8.0**

0.2 M MOPS, 50 mM sodium acetate, 10 mM EDTA with NaOH to pH 8.0 (stored at 4°C in the dark).

**1M Tris•HCl pH 8.0**

1 M Trizma base and HCl to pH 8.0.

**1X TBE**

90 mM Trizma base, 90 mM orthoboric acid, 2 mM EDTA.

**2.5 kb molecular ruler**

333 ng/μl of 2.5 kb molecular ruler in 1X TE, 1X Orange G loading dye.

**20X SSC**

3.0 M NaCl, 0.3 M sodium citrate.

**20X SSPE**

3.0 M NaCl, 0.2 M NaH<sub>2</sub>PO<sub>4</sub>, 20 mM EDTA, pH 8.0.

**3 M Sodium acetate pH 7.5**

3 M sodium acetate with NH<sub>3</sub> to pH 7.5.

**3.3X RNA sample buffer**

240 nM ethidium bromide, 2.3X MOPS pH 8.0, 5% (w/v) formaldehyde, 70 % (v/v) formamide (used on the day of making).

**5X DNA loading dye**

0.5% (w/v) SDS, 0.25% (w/v) xylene cyanol, 0.25% (w/v) bromophenol blue, 1.5% (w/v) Ficoll<sup>®</sup>400, in 3X TBE.

**5X Orange G loading dye**

0.06 % (w/v) Orange G, 50 % (v/v) glycerol in H<sub>2</sub>O.

**75% (v/v) Ethanol**

75% (v/v) absolute ethanol in H<sub>2</sub>O.

**80% (v/v) Ethanol**

80 % (v/v) absolute ethanol in H<sub>2</sub>O.

**Amplisize molecular ruler**

333 ng/μl of amplisize molecular ruler in 1X TE, 1X Orange G loading dye.

**Denaturing solution**

0.5 M NaOH, 1.5 M NaCl in H<sub>2</sub>O.

**Depurinating solution**

0.25 M HCl in H<sub>2</sub>O.

**Ethidium bromide**

Stock solution: 25 mM in H<sub>2</sub>O.

Working concentration: 500 nM.

**Genomic Southern blot hybridisation solution**

5X SSPE, 5X Denhardt's solution, 0.5% (w/v) SDS, 7% (w/v) dextran sulphate, 100 μg/ml salmon sperm DNA in H<sub>2</sub>O.

**Lysis buffer**

50 mM Tris•HCl pH 8.0, 100 mM EDTA pH 8.0, 0.5% (w/v) SDS in H<sub>2</sub>O.

**Neutralising solution**

1.5 M NaCl, 0.5M Trizma base and HCl to pH 7.5.

**PCR buffer**

45 mM Tris•HCl pH 8.8, 11 mM ammonium sulphate, 4.5 mM MgCl<sub>2</sub>, 6.7 mM 2-mercaptoethanol, 4.4 μM EDTA, 1 mM dATP, 1 mM dCTP, 1 mM dGTP, 1 mM dTTP and 113 μg/ml BSA.

**Phenol**

Phenol saturated in 10 mM Tris pH 8.0, 1 mM EDTA.

**Phenol:chloroform:isomayl alcohol, 25:24:1**

Phenol:chloroform:isomayl alcohol, 25:24:1, saturated in 10 mM Tris pH 8.0, 1 mM EDTA.

**Proteinase K**

Stock solution: 20 mg/ml proteinase K in filter sterilised 50 mM Tris•HCl pH 8.0.

**Squash blot hybridisation solution**

7% (w/v) SDS, 1M NaPO<sub>4</sub>, 2 mM EDTA.

**TAE**

40 mM Tris•acetate, 1 mM EDTA in H<sub>2</sub>O.

**TE**

10 mM Tris•HCl pH 8.0, 1 mM EDTA pH 8.0 in H<sub>2</sub>O.

**2.3.2. Bacterial solutions, media and antibiotics****Ampicillin**

Stock solution: 50 mg/ml in H<sub>2</sub>O (stored at -20°C).

Working concentration: 100 μg/ml.

**IPTG**

Stock solution: 100 mg/ml in H<sub>2</sub>O (stored at -20°C).

Working concentration: 10 μg/ml.

**Luria-Bertani (LB) medium**

3% (w/v) Bacto<sup>®</sup> tryptone, 0.5% (w/v) Bacto<sup>®</sup> yeast extract, 1% (w/v) NaCl.

LB agar contained 1.5% (w/v) agar.

**SOB medium**

2% (w/v) Bacto<sup>®</sup> tryptone, 0.5% (w/v) Bacto<sup>®</sup> yeast extract, 0.85 mM NaCl, 0.25 mM KCl pH 7.0 with NaOH. Sterilized 10 mM MgSO<sub>4</sub> added prior to use.

**SOC medium**

0.04% (w/v) glucose in SOB medium.

**X-gal**

Stock solution: 50 mg/ml in dimethylformamide (stored at -20°C).

Working concentration: 50 µg/ml.

**2.3.3. Tissue culture solutions, media and antibiotics**

Tissue culture media, serum, antibiotics and solutions were obtained from GibcoBRL Life Technologies or Sigma.

**Dulbecco's modified Eagle medium (DMEM)**

DMEM with 862 mg/l L-alanyl-L-glutamine (GlutaMax<sup>1</sup>), 4 mg/l pyridoxine•HCl, 4500 mg/l glucose, 110 mg/l sodium pyruvate.

**Trypsin/EDTA solution**

Stock solution: 5.0 g/l trypsin, 2.0 g/l EDTA, 8.5 g/l NaCl (stored at -20°C).

Working concentration: 0.5 g trypsin, 0.2 g EDTA•4Na/l in PBS.

**Penicillin and streptomycin solution**

Stock solution: 10,000 U/ml penicillin, 10,000 µg/ml streptomycin. Utilising penicillin G (sodium salt) and streptomycin sulphate: prepared in normal saline (stored at -20°C).

Working concentration: 100 U/ml penicillin, 100 µg/ml streptomycin.

**10X Dulbecco's phosphate buffered saline (PBS)**

8 g/l NaCl, 0.2 g/l KCl, 2 g/l KH<sub>2</sub>PO<sub>4</sub>, 2.16 g/l Na<sub>2</sub>HPO<sub>4</sub>•7H<sub>2</sub>O.

**Foetal bovine serum (FBS)**

Origin E.C. Virus and mycoplasma tested.

**2.4. Mouse tissue culture methods****2.4.1. Establishment of primary cell cultures**

*Dmt* transgenic mice were sacrificed by cervical dislocation, and the kidneys, eyes and lungs removed and kept on ice for up to 1 hour until processed. All mice were hemizygous for the transgene on a pure FVB/N genetic background. Depending on the tissue, the primary cultures

were established either by the explant technique (eye cultures) or enzymatic dissociation (lung and kidney).

#### **2.4.1.1. Eye cell cultures**

The eye balls were transferred to 100 mm dishes, rinsed and dissected in 1X phosphate buffered saline (PBS). The tissue was placed cell-side down on a 60 mm dish and minced into small fragments (~1 mm<sup>2</sup>). All pieces of tissue were flattened and dispersed on the dish. Culture medium was gently added and the culture was incubated at 37°C in a humidified 5% CO<sub>2</sub> atmosphere. The standard growth medium consisted of Dulbecco's modified Eagle medium with high glucose (DMEM) supplied with 10% foetal bovine serum (FBS), 100 U/ml of penicillin and 100 µg/ml of streptomycin. Once cells became 60-80% confluent they were subsequently subcultured at a 1:5 or 1:10 ratio during the first 5 passages and at a higher ratio (varying from 1:20 to 1:50) thereafter, as described in section 2.4.4.

#### **2.4.1.2. Lung and kidney cell cultures**

The enzymatic dissociation of kidney and lung tissues was essentially performed as described previously (Rybak and Murphy, 1998). The organs were transferred to 100 mm dishes containing 10 ml of sterile PBS and minced into 1 mm<sup>3</sup> cubes. The minced tissues were transferred into 15 ml Falcon tubes and washed twice with sterile PBS. The tissues were digested with 5 ml of trypsin/EDTA solution, in a 37°C humidified 5% CO<sub>2</sub> incubator for 30 minutes with limited shaking. After the incubation the pieces of tissue were allowed to settle down and the supernatants transferred into fresh 15 ml tubes. The cells in suspension were collected by centrifugation at 200 g for 5 minutes. The final pellets were resuspended in 5 ml of standard culture medium, plated on 25 cm<sup>2</sup> flasks and incubated at 37°C in a humidified 5% CO<sub>2</sub> atmosphere. The remaining portions of tissue were repeatedly digested with fresh trypsin for 6-12 times. After the first passage all the pellets derived from the same tissue were pooled together and an initial single culture was maintained for each tissue. Once the cells became ~80% confluent, they were subcultured at a 1:5 or 1:10 ratio during the first 5 passages, and at a higher ratio (varying from 1:20 to 1:50) thereafter, as described below in section 2.4.4.

### **2.4.2. Establishment of single cell-derived clones**

Clones derived from a single cell were established by seeding an average of 0.5 cells in each well of a 24 or 96-well cluster. Once the cells became confluent they were transferred into 25 cm<sup>2</sup> tissue culture flasks thereafter grown as described below (Sections 2.4.3 and 2.4.4).

### **2.4.3. Feeding cultured cells**

Cells were grown in a 5% CO<sub>2</sub> incubator at 37°C. Initially cells were cultured in 25 cm<sup>2</sup> flasks during the first 5-10 passages. All solutions (cell growth medium, PBS and trypsin/EDTA) were warmed to 37°C. The old medium (5 ml) was removed carefully, avoiding scratching the surface of the cells. Cells were washed gently with 3 ml of PBS, and 5 ml of fresh culture medium was added. The same standard protocol was adapted for cells growing in 6-well clusters: each well was washed with 2 ml of PBS and the cells grown in 3 ml of cultured medium.

### **2.4.4. Subculturing cultured cells**

Cells growing in 25 cm<sup>2</sup> flasks were washed in PBS as described before: 2 ml of trypsin/EDTA solution were added and the cells incubated at 37°C, 5% CO<sub>2</sub> for 5 minutes. The cells were observed under the microscope to check if they had rounded up and lifted off the surface of the culture flask. The flask was tapped gently against the bench to help the cells dissipate from the surface of the flask. If necessary, the cells were incubated for an extra two minutes and a cell scraper used to make sure that nearly every cell was floating. Trypsin digestion was stopped by the addition of 3 ml of fresh culture medium. The cell suspension was pipetted up and down repeatedly to dissociate cell clumps, and transferred into a 15 ml falcon tube. The cells were collected by centrifugation at 200 g for 5 minutes, the supernatant removed and the cell pellet resuspended in 5 ml of fresh culture medium. A variable amount of the cell suspension, depending on the split ratio (1:5 or 1:10), was collected and mixed with fresh growth medium, up to a final volume of 5 ml. For higher split ratios cells were not collected by centrifugation after trypsin digestion and prior to seeding. Instead, a small aliquot of the cell suspension (usually 100 µl or 125 µl for split ratios of 1:50 or 1:40, respectively) was mixed with 5 ml of fresh medium and plated in a fresh 25 cm<sup>2</sup> tissue culture flask. Cells were finally returned to the 37°C, 5% CO<sub>2</sub> incubator. At every passage the number of population doublings was determined based on the cell number.

When growing on 6-well clusters, cells were washed in 2 ml of PBS, digested with 0.5 ml of trypsin/EDTA solution, neutralised with 1.5 ml of standard growth medium and finally resuspended in 3 ml of fresh culture medium.

### **2.4.5. Measuring cell counts and determining population doubling times**

Following trypsin digestion and neutralisation with standard growth medium, cells were counted on a haemocytometer, using a phase contrast microscope. The haemocytometer was covered with the coverslip and a drop of cell suspension was dropped at the edge of the coverslip on both sides of the chamber. At least 100 cells were counted, and the number of cells/ml

calculated. The number of population doublings (PD) was determined, based on the cell number for two consecutive passages as follows:

$$PD = \log_2 \left( \frac{\text{cell number at passage } n+1}{\text{cell number at passage } n} \times \text{split ratio} \right)$$

Finally, the population doubling times were calculated by dividing the number of population doublings, by the time over which they had occurred:

$$PDT = PD / \text{time}$$

#### **2.4.6. Measuring cell viability**

To differentiate between live and dead cells in culture, acridine orange and ethidium bromide (both DNA intercalating agents) were used to determine cell viability. Acridine orange stains DNA bright green, whereas ethidium bromide stains DNA orange, but the latter is excluded by viable cells. As a result, living cells fluoresce green and dead cells fluoresce orange.

Cells were trypsinised and neutralised with standard growth medium. Equal volumes of cell suspension and ethidium bromide-acridine orange working solution (27.5  $\mu\text{M}$  ethidium bromide, 12.5  $\mu\text{M}$  acridine orange, 0.02% (v/v) ethanol in 1X PBS) were mixed together. The cells were subsequently counted on a haemocytometer under a fluorescence microscope with an excitation filter of 495 nm. Both live and dead cells were counted and viability was calculated as:

$$\text{Viability (\%)} = \frac{\text{Live cells}}{\text{Total cells counted}} \times 100$$

#### **2.4.7. Freezing cultured mouse cells in liquid nitrogen**

A single cell suspension was obtained following trypsin digestion as described before (section 2.4.4). Cells were precipitated by centrifugation at 200 g for 5 minutes, resuspended in fresh culture medium at a concentration of  $\sim 1.3\text{-}3 \times 10^6$  cells/ml and transferred into 2 ml cryovials. Dymethylsulphoxide (DMSO) was added to a final concentration of 10% (v/v), and then mixed well by gently inverting the tubes.

The vials were transferred into a freezing container with isopropanol at room temperature, and then cooled down to  $-70^\circ\text{C}$  for at least 4 hours. Finally the frozen samples were moved to liquid nitrogen ( $-190^\circ\text{C}$ ), where they were kept until needed.

### **2.4.8. Thawing cultured mouse cells**

Frozen cell samples stored in liquid nitrogen were quickly thawed by immersion of the cryovials in a water bath at 37°C. Five ml of fresh warm medium were added to the cells, and mixed gently by repeated pipetting. The cells were then collected by centrifugation at 200 g for 5 minutes, resuspended in fresh culture medium, and finally plated on a tissue culture flask or plate.

### **2.4.9. Immunocytochemical characterisation of cultured cell types**

Vimentin and cytokeratins were detected immunocytochemically in cultured mouse cells using mouse monoclonal antibodies raised against human proteins, which cross-react with mouse homologues. Semiconfluent cells growing on 8-well chamber slides were washed once with serum-free DMEM and fixed in 2% (w/v) paraformaldehyde in DMEM for 20 minutes at room temperature. The cells were washed once with PBS, left in 0.1 M glycine for 20 minutes and washed twice with PBS. Cells were permeabilised in 1% (v/v) Triton-X100 in PBS for 6 minutes, and washed twice with PBS. Monoclonal anti-vimentin (mouse IgM isotype; Sigma, catalogue number: V5255) or monoclonal anti-pan cytokeratins (mouse IgG1 isotype; Sigma, catalogue number: C1801) were applied at 1:200 in 0.01% (v/v) Triton-X100 and incubated over night at 4°C with gentle shaking. The cells were washed 4 times in PBS for five minutes. Both anti-mouse IgM-fluoresceine isothiocyanate isomer I (FITC) conjugate (Sigma, catalogue number: F9259) and anti-mouse IgG-tetramethylrhodamine isothiocyanate (TRITC) conjugate (Sigma, catalogue number: T7657) were applied in a 1:100 dilution in 3 mg/ml bovine serum albumin (BSA) in PBS with 0.01% (v/v) Triton-X100 and incubated for 2 hours at room temperature with gentle shaking. Cells were finally washed 4 times in PBS for 5 minutes and observed using fluorescence microscopy.

### **2.4.10. 5'-Bromo-3-deoxyuridine (BrdU) incorporation and detection assay**

In order to assess nuclear DNA synthesis and/or cell proliferation in cultured mouse cells, a 5'-bromo-3-deoxyuridine (BrdU) incorporation and detection assay was performed, following the manufacturer's protocol. This assay relies on the principle that BrdU can be incorporated into DNA molecules in place of thymidine, and later detected by monoclonal antibodies directed against BrdU. Cells that incorporate BrdU into their DNA can be detected using a fluorochrome-conjugated second antibody.

In brief, cells were grown on 8-well chamber slides until they reached ~60% confluency. The cell medium was removed and the cells washed once in PBS. The cells were incubated for 15 minutes, 60 minutes or 30 hours in standard growth medium containing 10 µM BrdU. The medium

was aspirated, the cells washed 3 times in “washing buffer”<sup>1</sup> for five minutes and fixed in 70% (v/v) ethanol, 15 mM glycine, pH 2.0 for 45 minutes at -20°C. The cells were washed 3 times in “washing buffer”, covered with anti-BrdU diluted 1:10 in “incubation buffer”<sup>1</sup> and incubated for 30 minutes at 37°C. Following the first antibody incubation, the cells were washed 3 times in “washing buffer” for 5 minutes, and incubated in anti-mouse-Ig-fluorescein, which had been diluted 1:10 in PBS, containing 10 mg/ml BSA. The cells were finally washed 3 times in “washing buffer”, and examined in a fluorescence microscope (excitation wavelength in the range 450-500 nm, and detection in the range of 515-565 nm).

## **2.5. Preparation, purification and analysis of DNA**

### **2.5.1. Preparation of plasmid DNA**

Qiagen kits were used in order to obtain high yields of good quality plasmid DNA. DNA was purified from 5-ml bacterial cultures by alkaline lysis followed by anion-exchange columns, according to the manufacturer’s protocol.

### **2.5.2. DNA extraction from mouse tissues**

Ten to 20 mg of tissue were minced and placed in a 1.5 ml screw top Eppendorf tube with 550 µl of lysis buffer and 15 µl of 20 mg/ml of proteinase K to a final concentration of 530 µg/ml of proteinase K. The tissue was incubated overnight at 60°C. The lysate was briefly centrifuged at 21,000 g to precipitate the debris. Two hundred and fifty µl of the supernatant were transferred into a fresh tube and an equal volume of phenol was added and mixed vigorously by vortexing for 5 seconds to emulsify the two phases. The mixture was centrifuged for 5 minutes at 21,000 g to separate the two phases. The upper phase was placed into a clean tube, and 250 µl of phenol:chloroform:isoamyl alcohol (25:24:1) were added and mixed vigorously by vortexing for 5 seconds. The upper phase was removed and transferred into a clean tube. In order to remove any remaining traces of phenol, 250 µl of chloroform were added and emulsified with the aqueous phase by vortexing for 5 seconds. The upper aqueous phase was removed and transferred into a clean tube, and 25 µl of 3 M sodium acetate pH 7.5 were added and mixed briefly by inverting the tube 5 to 10 times. The DNA was finally precipitated by the addition of 500 µl of ice-cold absolute ethanol, followed by incubation at -20°C for at least 1 hour, and centrifugation at 21,000 g in a bench top centrifuge. The supernatant was decanted off and the pellet rinsed with 1 ml of ice-cold 80% (v/v) ethanol, and precipitated again by centrifugation at 21,000 g in a top bench centrifuge. The DNA pellet was either air dried at room temperature for 30 to 60 minutes, or at 4°C overnight.

---

<sup>1</sup> Both “washing” and “incubation” buffers were supplied by the manufacturer

Once dried the DNA pellet was dissolved in 100 to 200  $\mu$ l of TE at 60°C for 30 minutes, or at 37°C overnight.

### **2.5.3. DNA extraction from cultured mouse cells**

Cell culture DNA samples were extracted using a Nucleon DNA extraction kit for blood and tissue culture, following the manufacturer's protocol. Briefly, cells were collected by centrifugation at 800 *g* for 5 minutes, resuspended in washing "Buffer A"<sup>2</sup> and left on ice for 5 minutes. Cells were collected by centrifugation at 1,300 *g* and subsequently lysed in "Reagent B"<sup>2</sup> by gently pipetting up and down, followed by a 30-minute incubation at 37°C. Proteins were precipitated by the addition of 5M sodium perchlorate and subsequently extracted by the addition of chloroform, Nucleon resin and centrifugation at 370 *g* for 2 minutes. The upper aqueous phase was removed and transferred into a fresh 1.5 ml Eppendorf tube, without disturbing the resin layer or the chloroform phase. An extra centrifugation step at 5,000 *g* for 2 minutes was performed to pellet down any resin that might have been carried over. The supernatant was collected and transferred into a fresh 1.5 ml Eppendorf tube. DNA was precipitated by the addition of ice-cold absolute ethanol and centrifugation at 21,000 *g* for 10 minutes. The supernatant was discarded and the DNA pellet washed in ice-cold 80% (v/v) ethanol and either air-dried at room temperature for 30 minutes, or at 4°C overnight. Once dried, the DNA pellet was dissolved in 100  $\mu$ l of TE at 60°C for 30-60 minutes, or at 37°C overnight, and finally stored at -20°C.

### **2.5.4. Determination of DNA concentration**

When necessary, the concentration and purity of DNA in aqueous solution were estimated by measuring the UV absorbance of the solution at wavelengths ranging from 200-300 nm. Pure double-stranded DNA solutions have an absorbance maximum at 260 nm, at which an optical density of 1 corresponds to 50 mg/ml of DNA in the solution. The purity of the DNA samples was estimated by comparing the ratio of the OD at 260 nm (OD<sub>260</sub>) to 280 nm (OD<sub>280</sub>), at which wavelength proteins have an absorbance maximum. An OD<sub>260</sub>/OD<sub>280</sub> ratio of 1.8 or greater is taken as an acceptable level of purity. DNA samples were diluted 1:100 in H<sub>2</sub>O and the spectrophotometer baseline corrected with H<sub>2</sub>O. The OD was measured and the concentration calculated as 100 x 50 x OD<sub>260</sub> mg/ml.

Alternatively, DNA concentrations were determined by electrophoresis on agarose gels, followed by densitometry analysis against low molecular weight DNA mass ladder (GibcoBRL, Life Technologies), using Kodak Digital Science 1D software (Kodak).

---

<sup>2</sup> Both "Buffer A" and "Reagent B" were supplied by the manufacturer.

### **2.5.5. Polymerase chain reaction (PCR)**

All PCR amplifications were carried out in either a Biometra Uno thermal cycler or a Biometra T3 thermal cycler, in either Costar 96 well plates or Anachem thin walled 200  $\mu\text{l}$  tubes. Each reaction was overlaid with white light mineral oil and then sealed. The thermal cycler lid was preheated to 105°C.

Except for small pool PCR analysis (section 2.7.2), standard PCR amplifications were performed with 1  $\mu\text{M}$  of each primer, 1X PCR buffer and 0.5 U *Taq* DNA polymerase in a total volume of 10  $\mu\text{l}$ . The annealing temperature varied according to the primers selected to carry out the amplification, but in general fell within a range from 60°C to 68°C. Reactions were thermal cycled under the following conditions:

20-35 cycles:	96°C for 45 seconds
	60-68°C for 45 seconds
	70°C for 3 minutes
1 step:	60-68°C for 1 minute
1 step:	70°C for 10 minutes

### **2.5.6. DNA cloning techniques**

#### **2.5.6.1. Restriction endonuclease digestion of plasmid DNA**

Restriction digests were carried out to generate gene fragments for checking the inserts of recombinant plasmids containing foreign DNA.

1-5  $\mu\text{g}$  of purified plasmid DNA, 5-10 units of restriction endonuclease, 1X recommended restriction endonuclease reaction buffer in a total of 10  $\mu\text{l}$  with  $\text{H}_2\text{O}$  were incubated at 37°C for 90-120 minutes.

#### **2.5.6.2. DNA ligation**

Bacteriophage T4 DNA ligase was used to catalyse the formation of phosphodiester bonds between the 3'-hydroxyl groups and the 5'-phosphate groups of the DNA inserts and vectors. A recombinant plasmid with 2 single-stranded nicks results owing to the removal of the 5'-phosphate groups from the digested vector. The nicks are repaired after the recombinant plasmids have been introduced into competent bacteria.

Ligations of vector and insert DNA were incubated at 4°C overnight with 1 U of bacteriophage T4 DNA ligase and 1X T4 DNA ligase buffer in a total volume of 10  $\mu\text{l}$ . Generally, DNA molecules were ligated at vector:insert molar ratio of 1:3. Control ligations, with one component missing, were also incubated at 4°C overnight.

### 2.5.6.3. Transformation of competent bacterial cells

Fifty  $\mu\text{l}$  aliquots of competent cells were thawed on ice for 10 minutes. One  $\mu\text{l}$  of the ligation mixture or 10 ng of control plasmid were added and after gentle mixing, the bacteria were incubated on ice for 30 minutes. The cells were heat-shocked at  $42^{\circ}\text{C}$  for 45 seconds and cooled on ice for 2 minutes. Pre-warmed SOC medium (450  $\mu\text{l}$ ) was added to the cells and the samples incubated at  $37^{\circ}\text{C}$  for 1 hour. The transformed cells were split into two aliquots of 50  $\mu\text{l}$  and 450  $\mu\text{l}$ , which were subsequently spread on a Luria-Bertani (LB) plate containing 100  $\mu\text{g/ml}$  ampicillin, 50  $\mu\text{g/ml}$  5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactoside (X-gal) and 10  $\mu\text{g/ml}$  isopropylthio- $\beta$ -D-galactoside (IPTG), and allowed to dry for up to 30 minutes, before overnight incubation at  $37^{\circ}\text{C}$ . Ampicillin, X-gal and IPTG are used to positively select transformed bacteria, which are resistant to the antibiotic ampicillin, and are white because of the insertional inactivation of the *lacZ* gene of the recombinant plasmid. As a result  $\beta$ -galactosidase is not synthesised and the bacterial colonies are white rather than blue.

### 2.5.6.4. Generation of plasmid stocks

Stocks of bacteria transformed with recombinant plasmids were prepared for long-term storage. A colony from an agar plate was used to inoculate 5 ml of LB medium, containing 100  $\mu\text{g/ml}$  of ampicillin and incubated at  $37^{\circ}\text{C}$  overnight. An aliquot of overnight culture (usually 500  $\mu\text{l}$ ) was mixed with an equal volume of 2% (w/v) peptone, 40% (v/v) glycerol, by repeated inversion. The stocks were stored at  $-70^{\circ}\text{C}$  until needed.

## 2.5.7. Gel electrophoresis

### 2.5.7.1. Agarose gel electrophoresis

DNA molecules were separated according to their size by agarose gel electrophoresis. Solutions of 0.8-2% (w/v) of agarose in 0.5X TBE were prepared in a microwave. When the solution had cooled to approximately  $50^{\circ}\text{C}$ , ethidium bromide was added to a final concentration of 500 nM and the gel was cast in a horizontal tray. DNA samples were mixed with 1X DNA loading dye and electrophoresed in 0.5X TBE for periods of time varying from 2-16 hours to overnight at  $\sim 1$ -4 volts/cm. An aliquot of a DNA size ladder (usually 100-200 ng) was also electrophoresed and used as a size marker. Separated DNA samples were visualised using an UV transilluminator (wavelength 254 nm) and photographed.

### 2.5.7.2. Non-denaturing polyacrylamide gel electrophoresis

Non-denaturing polyacrylamide gel electrophoresis (PAGE) was carried out in a BioRad Protean II gel apparatus, which included 15 x 14 cm separating gels. 8% (w/v) non-denaturing acrylamide/bis (29:1) gels, containing 10% (v/v) glycerol were prepared in 1X TBE. The

polymerising agents *NNN'*-tetramethylethylenediamine (TEMED) and ammonium persulphate (APS) were added to a final concentration of 0.625% (v/v) and 0.125% (w/v), respectively. Once the gel polymerised, 100 V were applied for 90 minutes with continuous re-circulation of the running buffer (1X TBE) at 4°C.

Each DNA sample was mixed with an equal volume of 5X loading dye and electrophoresed at 140V for 16 hours at 4°C with buffer re-circulation. The gel was stained with 500 nM ethidium bromide in 1X TBE for 20 minutes at room temperature. Separated DNA samples were visualised using an UV transilluminator (wavelength 254 nm) and photographed.

### **2.5.7.3. DNA extraction from non-denaturing polyacrylamide gels**

Bands corresponding to the DNA samples of interest were cut out from the polyacrylamide gel under an UV transilluminator. DNA was eluted from the gel fragments by simple diffusion, in 500 µl of 1X TE, at 37°C for 48 hours. Given the low efficiency of the elution process, very low amounts of DNA were expected to be recovered from the gel bands, and therefore linear polyacrylamide was used as a DNA carrier to precipitate gel-purified DNA samples.

Linear polyacrylamide stock solution was prepared as follows. A 5% (w/v) acrylamide solution in 40 mM Tris•HCl, 20 mM sodium acetate, 1 mM EDTA, pH 7.8 was used in the polymerisation reaction, which was initiated by the addition of 1/100 volume of 10% (w/v) ammonium persulphate, and 1/1000 volume of TEMED. Acrylamide was allowed to polymerise for 30 minutes. When the solution became viscous the polymer was precipitated with 2.5 volumes of ethanol, centrifuged at 21,000 *g* for 10 minutes and redissolved in 20 volumes of water, to obtain a 0.25% (w/v) linear polyacrylamide solution, which was stored at 4°C.

Gel-purified DNA was precipitated by the addition of 50 µl of 3M sodium acetate, pH 5.3, 25 µg of linear polyacrylamide and 1.5 ml of ice-cold absolute ethanol. The mixture was left at -70°C for 30 minutes and then centrifuged at 21,000 *g* for 10 minutes. The supernatant was carefully removed, and the pellet washed in ice-cold 70% (v/v) ethanol. DNA samples were air-dried, resuspended in 30-50 µl of 1X TE and stored at -20°C.

## **2.5.8. Southern blotting**

### **2.5.8.1. Genomic DNA transfer from agarose gels onto a nylon membrane by Southern blotting**

Gels were rinsed in deionised water, incubated in depurinating solution for 10 minutes, denaturing solution for 30 minutes and neutralising solution for 30 minutes and rinsed in deionised water between incubations. Incubations were performed with gentle shaking. The capillary blot was assembled using Hybond<sup>TM</sup>-N<sup>+</sup> nylon membrane and 20X SSCE. Sheets of Hybond<sup>TM</sup>-N<sup>+</sup> were rinsed first in deionised water and then in 20X SSPE. All sheets of gel blotting paper were rinsed in 20X SSPE before assembly, and care was taken to eliminate any air bubbles. Gels were blotted for

16 hours and then the membrane was baked at 80°C for 2 hours and finally crosslinked with 1,200 J/m<sup>2</sup>.

#### **2.5.8.2. DNA transfer from agarose or polyacrylamide gels onto a nylon membrane by Southern “squash” blotting**

Once the gels had run sufficient distance, a scalpel and a stainless steel ruler were used to cut off excess gel not required. Gels were rinsed in deionised water, incubated in depurinating solution for 10 minutes, denaturing solution for 30 minutes and neutralising solution for 30 minutes and rinsed in deionised water between incubations. Incubations were performed with gentle shaking.

A piece of Magna nylon membrane the same size as the gel was wet in deionised water first, and then in neutralising solution. A piece of Saran Wrap, somewhat larger in size than the gel, was placed on the bench, with a piece of gel blotting paper, the same size as the gel and wet in neutralising solution, layered on top. The gel was placed on the Saran Wrap, and the membrane layered on top of the gel. Any air bubbles trapped under the membrane were carefully removed. Two sheets of gel blotting paper were rinsed in neutralising solution and layered onto the membrane. The blot was topped with a thick layer of paper towels, a glass plate and a weight (500 g for a small gel, 1 kg for a large gel). The DNA was transferred from the gel onto the membrane by capillary action for 3 up to 16 hours. The blot was dismantled in reverse order and the membrane placed on a piece of dry gel blotting paper, the DNA side up. The membrane was baked for at least 20 minutes in an 80°C oven and the DNA fixed to the membrane by exposure to 1,200 J/m<sup>2</sup> in a UV crosslinker.

#### **2.5.9. Preparation of radiolabelled double-stranded probes**

Double-stranded DNA probes were radiolabelled using  $\alpha$ -<sup>32</sup>P-labelled 2'-deoxycytidine-5'-triphosphate ([ $\alpha$ -<sup>32</sup>P]dCTP) and the Ready-to-go kit, following the manufacturer's instructions. Briefly, 30 ng of DNA double-stranded probe, 5 ng of DNA size marker and H<sub>2</sub>O to a final volume of 45  $\mu$ l were denatured at 100°C for 5 minutes and quenched on ice for 3 minutes. The lyophilised reaction mix was resuspended in the DNA solution. Five  $\mu$ l of [ $\alpha$ -<sup>32</sup>P]dCTP (3,000 Ci/nmol) were added and the mixture incubated at 37°C for 1 hour. At the end of the reaction time, 250  $\mu$ l of H<sub>2</sub>O were added to stop the reaction.

Double-stranded DNA probes radiolabelled for genomic Southern blot hybridisation, and northern blot hybridisation, were purified using a 1.0 ml Sephadex<sup>®</sup> G-50 column, prepared in a 1.0 ml plastic syringe. The Sephadex<sup>®</sup> was compacted by centrifugation at 1,800 g for 2 minutes. The probe was added to the column and centrifuged at 2,000 g for 3 minutes. The eluate was collected and the percentage of incorporation calculated by taking counts per minute (cpm)

readings before and after purification. Probes used for the detection of PCR products were not purified, since a low background was expected.

## **2.5.10. Southern hybridisation**

### **2.5.10.1. Hybridisation of genomic DNA transferred onto a nylon membrane**

Filters to be hybridised were placed in a hybridisation bottle with 20 ml of genomic Southern blot hybridisation solution and rotated at 65°C for 16 hours. The purified radiolabelled probe was denatured at 100°C for 5 minutes and quenched for 2 minutes on ice before being added to the bottle. Hybridisation was carried out at 65°C overnight in a rotating hybridisation oven.

Filters were washed flat with gentle shaking in 0.1% (w/v) sodium dodecyl sulphate (SDS), sequentially to the following stringencies:

2X SSPE	room temperature	20 minutes
2X SSPE	65°C	20 minutes
1X SSPE	65°C	20 minutes
0.1X SSPE	65°C	20 minutes

At each stage the filters were monitored with a Geiger counter, and washing was stopped when the reading was below 10 counts per second (cps). Filters were sealed in a plastic bag whilst still damp, and exposed to X-ray film in the presence of an intensifying screen at -70°C. Autoradiographs were developed after a two- to three-week exposure.

### **2.5.10.2. Hybridisation of PCR products and plasmid DNA transferred onto a nylon membrane**

Filters to be hybridised were placed in a hybridisation bottle and rotated at 65°C for 30 minutes with 10 ml of squash blot hybridisation solution. The pre-hybridisation step was repeated twice. The radiolabelled probe was denatured at 100°C for 5 minutes and quenched on ice for two minutes before being added to the bottle, which contained 10 ml of fresh squash blot hybridisation solution. Hybridisation was performed at 65°C overnight in a rotating hybridisation oven.

Following hybridisation, the filters were briefly washed inside the bottle in 0.2% (w/v) SDS, 0.2X SSPE at room temperature to remove the excess of probe and free [ $\alpha$ -<sup>32</sup>P]dCTP. The filters were then washed twice in the same high stringency washing solution for 30 minutes at 65°C. Finally the filters were transferred onto a flat tray and washed by gently shaking them at room temperature in the same solution. The filters were baked at 80°C until dry, and exposed to X-ray film at room temperature. Autoradiographs were developed after an exposure time varying from 4 hours to 3 days.

## **2.6. Preparation, purification and analysis of RNA**

### **2.6.1. RNA extraction from cultured mouse cells**

Since RNA is vulnerable to degradation from nucleases, to optimise the successful handling of RNA, the following precautions were taken. RNA work was restricted to a dedicated bench. Dedicated equipment and stock reagents were used wherever possible and gloves were worn at all times. All solutions required for RNA related techniques were made using diethylpyrocarbonate (DEPC) treated water. DEPC was added to a concentration 0.1% (v/v) to deionised water, mixed and allowed to stand overnight before the inactivation of the DEPC by autoclaving. All plasticware was soaked in active DEPC-treated water before autoclaving, all glassware was soaked in active DEPC water before baking.

Mouse cells were grown on a 75 cm<sup>2</sup> tissue culture flask until they became confluent. Approximately  $6 \times 10^6$  cells were collected after trypsin digestion (see section 2.4.4), washed once in 1X PBS and lysed in 1 ml of Tri Reagent™ by repeated pipetting. The lysate was allowed to stand for 5 minutes, to ensure complete dissociation of nucleoprotein complexes. Following cell lysis, 200 µl of chloroform were added, the mixture was shaken vigorously for 15 seconds and allowed to stand for 15 minutes at room temperature. The mixture was then separated into three phases by centrifugation at 12,000 g for 15 minutes at 4°C. The aqueous phase was transferred into a fresh tube, 500 µl of isopropanol were added and the mixture was allowed to stand for 10 minutes at room temperature. The RNA was precipitated by centrifugation at 12,000 g for 10 minutes at 4°C. The supernatant was removed, the RNA pellet washed in 1 ml of 75% (v/v) ethanol and collected by centrifugation at 7,500 g for 5 minutes at 4°C. The RNA pellet was let to air-dry for 10-30 minutes and finally resuspended in 40 µl of 1X TE, 0.1% (w/v) SDS. Repeated pipetting at 60°C for 15 minutes facilitated RNA solubilisation.

### **2.6.2. Determination of RNA concentration**

The concentration and purity of RNA in aqueous solution was estimated by measuring the UV absorbance of the solution at wavelengths ranging from 200-300 nm. Pure RNA solutions have an absorbance maximum at 260 nm, at which a solution containing 40 mg/ml of RNA gives rise to an absorbance of 1. The purity of the DNA samples was estimated by comparing the ratio of the OD at 260 nm to 280 nm, at which wavelength proteins have an absorbance maximum. An OD<sub>260</sub>/OD<sub>280</sub> ratio of 1.65 or greater is taken as an acceptable level of purity. RNA samples were diluted 1:80 in H<sub>2</sub>O and the spectrophotometer baseline corrected with H<sub>2</sub>O. The OD was measured and the concentration calculated as  $80 \times 40 \times \text{OD}_{260}$  mg/ml.

### **2.6.3. RNA visualisation by agarose gel electrophoresis**

In order to assess RNA integrity and quality, RNA samples were electrophoresed through a 1% (w/v) agarose gel containing 0.1% (w/v) SDS, 500 nM ethidium bromide, and prepared in 1X TAE. Five µg of RNA were diluted in DNA loading buffer in a final volume of 10 µl, heated at 65°C for 5 minutes and immediately cooled on ice. The samples were electrophoresed at 100 V for 1 hour. 100-200 ng of 1 kb<sup>+</sup> DNA size marker was also electrophoresed. Separated RNA molecules were visualised using a UV transilluminator (wavelength 254 nm) and photographed.

### **2.6.4. Northern blotting and hybridisation**

#### **2.6.4.1. RNA electrophoresis in denaturing agarose gels**

A solution containing 1.3% (w/v) of agarose in DEPC treated water was prepared and heated in a microwave. Shortly after boiling 10X 3-(*N*-morpholino)-propanesulphonic acid (MOPS) buffer H 8.0 was added to a final concentration of 1X, and 40% (w/v) formaldehyde was added to a final concentration of 4% (w/v). The solution was allowed to cool to approximately 50°C and only then was the gel cast on a horizontal tray. Fifteen µg RNA aliquots in DEPC treated water were made up to a final volume of 20 µl in 1X RNA sample buffer and 1X DNA loading dye. RNA samples were subsequently heated at 65°C for 5 minutes, immediately cooled on ice and finally electrophoresed in 10X MOPS pH 7.0, at 150V, for 4 hours in a fume hood. An aliquot of RNA size ladder (100-200 ng) was also electrophoresed and used as an RNA size marker. Separated RNA samples were visualised using a UV transilluminator (wavelength 254 nm). The integrity of separated RNA molecules, in particular the 28S and 18S ribosomal RNA molecules, was checked and the gel photographed beside a fluorescent ruler.

#### **2.6.4.2. RNA transfer from denaturing agarose gels onto a nylon membrane by northern blotting**

The denaturing gels were rinsed in deionised water, incubated in 6X SSC for 30 minutes. The RNA was transferred onto a biodyne B nylon membrane by capillary action, in 10X SSC. Sheets of biodyne B nylon membrane were rinsed first in deionised water and then in 10X SSC. All sheets of gel blotting paper were rinsed in 10X SSC prior to blot assembly, and care was taken to eliminate any air bubbles. Gels were blotted for 16 hours and then the membrane was baked at 80°C for 2 hours and finally crosslinked with 1,200 J/m<sup>2</sup>.

#### **2.6.4.3. Hybridisation of RNA transferred onto a nylon membrane**

The membranes to be hybridised were placed in a hybridisation bottle with 5 ml of ExpressHyb hybridisation solution and rotated at 68°C for 1 hour. The purified radiolabelled DNA

probes were denatured at 100°C for 5 minutes and quenched on ice for 2 minutes, before being added to the bottle. Hybridisation was carried out at 68°C for 2 hours in a rotating hybridisation oven. The filters were washed flat with gentle shaking in 0.1X SSC, 0.1% (w/v) SDS, at 65°C for 30 minutes. After each wash the filters were monitored with a Geiger counter, and washing was stopped if the background reading was below 10 cps. Filters were sealed in a plastic bag whilst still damp, and exposed to X-ray film in the presence of an intensifying screen at -70°C or at room temperature. Autoradiographs were developed after an exposure time varying from 1 hour to 3 days.

## 2.7. Small pool PCR analysis

Small pool polymerase chain reaction (SP-PCR) allows for the detection of infrequent heterogeneous mutant alleles by limiting the number of DNA molecules to be amplified (Jeffreys *et al.*, 1994; Monckton *et al.*, 1995).

### 2.7.1. Restriction endonuclease digestion of mouse genomic DNA

Prior to PCR amplification mouse genomic DNA was digested with *Hind*III restriction endonuclease. Twenty µl of genomic DNA stock solution (approximately 20-100 µg of genomic DNA) were digested with 10 U of *Hind*III in the presence of 10 mM spermidine, in a final volume of 100 µl, for 2-16 hours at 37°C. Following digestion the enzyme was inactivated at 65°C for 10 minutes, and the digested DNA samples stored at -20°C.

### 2.7.2. Small pool PCR amplification

SP-PCR analyses (Jeffreys *et al.*, 1994; Monckton *et al.*, 1995) were performed to assess the levels of trinucleotide repeat instability in a particular DNA sample.

*Hind*III-digested mouse genomic DNA was serially diluted in 1X TE, containing 0.1 µM of carrier primer. As a general rule, the forward oligonucleotide primer used in subsequent PCR amplifications, was used as a carrier primer to prepare the dilution buffer. The dilutions usually fell in the following range: 1/10 to 1/100,000, depending on the DNA concentration in the original stock solution. Four different dilutions were prepared for each DNA sample, and 8 independent PCR amplifications were set up for each individual dilution.

The amplification of 0.5 µl of each DNA solution was carried out in a final volume of 7 µl, with 0.175 U of *Taq* DNA polymerase, 0.2 µM of each primer, in 1X PCR buffer. Except where otherwise stated, the annealing temperature was set at 68°C. The reactions were cycled in a Biometra Uno thermocycler as follows:

28 cycles:	96°C for 45 seconds
	68°C for 45 seconds
	70°C for 3 minutes
1 step:	68°C for 1 minute
1 step:	70°C for 10 minutes

SP-PCR amplifications were usually carried out using PCR primers DM-C and DM-BR, unless otherwise stated.

### **2.7.3. Gel electrophoresis of SP-PCR products**

Following PCR amplification, 3  $\mu$ l of DNA loading dye were added to each 7  $\mu$ l reaction, and 5  $\mu$ l of the final mixture was loaded on a 1.25% (w/v) agarose gel (20 x 50 cm) in 0.5X TBE. The initial power of 300 V was applied for 30 minutes to ensure minimal DNA diffusion from the wells into the buffer. The PCR products were subsequently resolved at 160-180 V for 16 hours at 4°C in 0.5X TBE. DNA ladders were also electrophoresed, and used as DNA size markers.

### **2.7.4. DNA transfer from the gel onto nylon membranes by Southern “squash” blot**

Once the gels had run sufficient distance, the DNA was transferred from the gel onto a magna nylon membrane and hybridised with a double stranded DNA probe (DM56) comprising a CAG•CTG repetitive sequence (see section 2.2.6) as described in sections 2.5.8.2 and 2.5.10.2, respectively). The PCR products were sized using Kodak Digital Science 1D software (Kodak).

### **2.7.5. Single molecule SP-PCR analysis**

High DNA input SP-PCR amplifications are useful as an indication of the general pattern of repeat instability and also to detect rare expansion or deletion events. However, the amplification of a high amount of DNA per reaction does not allow the precise quantification of the repeat number distribution in a given sample. A precise method of quantitative analysis was required to plot accurate repeat number frequencies for a population of cells.

Preliminary SP-PCR amplifications at low DNA concentrations were used to determine more accurately the number of molecules amplified in each reaction, according to the Poisson distribution. As the DNA concentrations have been derived by serial dilution, the bands observed in a single lane (hereafter named positive lane) may result from the amplification of multiple original template molecules. The use of positive lanes to calculate the average number of molecules

amplified in each reaction leads, therefore, to an underestimation the DNA concentration. Instead, the frequency of blank lanes (lanes lacking amplification products) appears to be an adequate alternative for the calculation of the DNA concentration in a given sample, and was determined as follows:

$$p(0) = \frac{\text{total number of lanes} - \text{number of positive lanes}}{\text{total number of lanes}}$$

The average number of molecules amplified per reaction at the chosen DNA concentration to carry out this calculation, was derived from the Poisson distribution:

$$\text{Average number of molecules per reaction} = -\ln p(0)$$

From these calculations it was possible to carry out SP-PCR amplification of an average of one transgene molecule per reaction. Around 100 transgene molecules were usually taken as the minimum number that should be sized for each time point, to obtain a repeat distribution representative of the whole population of cells.

### 3. Establishment and characterisation of a mouse cell culture model system of trinucleotide repeat instability

#### 3.1. Introduction

The expansion of unstable trinucleotide sequences in the human genome is the primary genetic defect associated with a specific class of complex human disorders, including myotonic dystrophy type 1 (DM1), Huntington disease (HD), fragile X syndrome (FRAXA), Friedreich ataxia (FRDA) and an ever-increasing number of spinocerebellar ataxias (SCAs) (Cummings and Zoghbi, 2000a; Cummings and Zoghbi, 2000b). Most of these diseases, such as DM1, HD and the SCAs are caused by the expansion of CAG•CTG repeats. In addition to a common genetic defect, these conditions also share an unusual pattern of inheritance, known as genetic anticipation, that is, a decreasing age of onset and the worsening of symptoms through successive generations (Harper *et al.*, 1992). At the molecular level, clinical anticipation relies on the expansion-biased germline instability of expanded trinucleotide repeats as they are transmitted from one generation to the next. Moreover, the repeats are also somatically unstable, with differences in repeat mutation profiles being commonly observed, not only between different tissues from the same individual, but also within the same tissue. It is generally assumed that variability arises through DNA replication slippage during cell division (Richards and Sutherland, 1994; Wells *et al.*, 1998). Such a mechanism might be facilitated by the propensity of these sequences to adopt non-orthodox non-B-DNA secondary structures, such as slipped-stranded DNA (S-DNA) (Pearson and Sinden, 1996). Alternative DNA conformations could induce polymerase stalling and/or stabilise replication intermediates. Attractive as this suggestion is, there is as yet no direct evidence from a mammalian system to support the existence of such structures *in vivo*, nor the hypothesis that mutations arise during DNA replication. Indeed, there are no obvious correlations between human and mouse tissues in which the repeats are more prone to change and rates of cell turnover *in vivo* (Anvret *et al.*, 1993; Ashizawa *et al.*, 1993; Fortune *et al.*, 2000; Hashida *et al.*, 2001; Kennedy and Shelbourne, 2000; Lia *et al.*, 1998; Monckton *et al.*, 1995; Seznec *et al.*, 2000; Thornton *et al.*, 1994). However, it should be considered that such analyses have been primarily performed on whole tissues or organs comprised of multiple cell types, most likely to have very differing dynamics in terms of both cell turnover and repeat metabolism. Such a complex scenario confounds attempts to establish simple correlations and may have masked subtle but valid relationships.

In order to gain greater insight into the molecular mechanisms underlying triplet repeat dynamics, trinucleotide repetitive tracts have been cloned into simple model organisms, such as *E. coli* (Bacolla *et al.*, 1998) and *S. cerevisiae* (Jinks-Roberston *et al.*, 1998). These systems have proved useful and indeed revealed some important insights into the factors affecting repeat

dynamics, such as orientation with respect to replication origins, transcription status of the repeat and DNA repair and/or recombination gene mutations. However, the repeats are inherently biased towards contraction in such systems, in complete contrast to the predominantly expansion-biased behaviour observed at most loci in humans (Freudenreich *et al.*, 1997; Kang *et al.*, 1995b; Schmidt *et al.*, 2000; Schweitzer and Livingston, 1997; Wells *et al.*, 1998)

So far the analysis of repeat stability in cell cultures of patient-derived cells has yielded mixed results. The DM1 repeat continues to expand *in vitro* in primary dura mater or skeletal muscle cell cultures, derived from a DM1 foetuses, and the dynamics profile of the expansion *in vitro* closely resembled a sigmoid function of culture time (Wörhle *et al.*, 1995). In contrast, Epstein-Barr virus (EBV)-transformed lymphoblastoid cell lines (LBCLs) derived from DM1 patients reveal an unusual pattern of repeat variability, with two types of mutations being detected. In addition to the frequent small length change mutations, which are biased towards expansion, rare, but very large, deletion mutants are also observed at a frequency exceeding that detected *in vivo* (Ashizawa *et al.*, 1996; Khajavi *et al.*, 2001). These results suggest that the EBV-transformation process results in altered cellular DNA metabolism. Such an effect might be expected, given that some EBV proteins alter the transcriptional activity of cellular genes involved in DNA processing (Kuhn-Hallek *et al.*, 1995). It has also been reported that expanded GAA•TTC repeats at the *FRDA* locus in EBV-transformed LBCLs display contractions and expansions of similar magnitudes and frequencies (Bidichandani *et al.*, 1999). Thus, it appears that although EBV-transformed LBCLs may be useful for modelling some aspects of repeat metabolism, their failure to accurately recreate the *in vivo* dynamics limits their overall utility. Moreover, although primary human cell cultures may be a good model, not only their availability from individuals with rare inherited disorders is severely limited, but also further dissection of the factors affecting instability, using patient-derived tissue samples, is complicated by confounding influences of patient age, progenitor allele length, tissue type, and genetic background, *et cetera*, on mutation rates and directions.

In order to create additional model systems in which repeat biology may be assessed *in vivo* in a mammalian system, transgenic mice containing unstable expanded CTG•CAG arrays have been created (Gourdon *et al.*, 1997; La Spada *et al.*, 1998; Mangiarini *et al.*, 1997; Monckton *et al.*, 1997; Sato *et al.*, 1999; Shelbourne *et al.*, 1999; van Den Broek *et al.*, 2002; Wheeler *et al.*, 1999b). To examine the intrinsic stability of CAG•CTG, five transgenic mouse lines carrying a portion of the *DMPK* 3'-UTR with 162 CTG repeats, derived from the human *DM1* locus, were previously generated by Monckton *et al.* (1997). One line in particular, *Dmt-D*, which comprises a single copy of the construct, has been shown to reproduce the dramatic tissue-specific, age-dependent and expansion-biased repeat instability associated with somatic mosaicism in DM1 patients (Fortune *et al.*, 2000). Similar mutational dynamics have been reported in transgenic models carrying large 45 kb fragments of the human *DM1* locus, incorporating ~300 repeats (Seznec *et al.*, 2000). In contrast, the *Dmt-E* mouse line, also carrying a single copy of the *Dmt162* transgene, yet inserted elsewhere in the mouse genome, does not exhibit detectable levels of trinucleotide repeat instability in the soma (Fortune *et al.*, 2000). A role for the flanking DNA

sequences as a modifier of repeat instability was previously suggested, and striking correlations between mutation rate and flanking GC content and proximity to CpG islands have been found in humans (Brock *et al.*, 1999). Taken together, these observations strongly suggest that global *cis*-acting mutational modifiers affect the relative stability of expanded CAG•CTG repeats.

Transgenic mouse models that mimic the trinucleotide somatic instability observed in patients represent a powerful tool to resolve the molecular mechanisms regulating repeat dynamics in a mammalian environment. The development of a cell culture system from such animals, which faithfully reproduces the *in vivo* trinucleotide dynamics *in vitro*, may create new avenues to investigate the multiple factors affecting the metabolism of repetitive sequences under controlled conditions. To this end, cell cultures could be established from tissue samples harvested from *Dmt-D* transgenic mice in order to monitor the CTG•CAG repeat size variability over extensive time periods, aiming to clarify the effect of a large number of cell divisions on trinucleotide repeat stability. Such a model system should provide new insights into the cellular metabolism of the simple CAG•CTG tandem repeats associated with inherited human disease.

## 3.2. Results

### 3.2.1. Mouse genotyping for the *Dmt162* transgene

With the aim of establishing a mouse cell culture system capable of mimicking the complex dynamics of expanded trinucleotide repeats, *Dmt-D* mice (Monckton *et al.*, 1997) have been selected as a reliable and reproducible source of tissue samples. These mice carry a single copy of the *Dmt162* transgene, randomly integrated in their genome, and faithfully recreate the age-dependent, tissue-specific, expansion-biased somatic mosaicism detected in human patients (Fortune *et al.*, 2000).

*Dmt* hemizygous mice<sup>3</sup> were identified by PCR amplification of transgenic sequences using 10-100 ng of tail DNA as template and different set of oligonucleotide primers in 10 µl reactions (Figure 3.1, Table 3.1). Primers mUSF-A and mUSF-BR were used to amplify a 1019 bp fragment within the *mouse upstream stimulatory factor (mUSF) 2* gene. The generation of a PCR product in this reaction confirmed the presence of mouse DNA in the sample. Transgene-specific primers DM-R and DM-QR amplify a 175 bp transgenic sequence, which maps 3' to the CAG•CTG repetitive tract, establishing the presence of the *Dmt162* transgene within the mouse genome. Transgene-specific primers DM-C and DM-BR amplify across the transgenic CAG•CTG tract, giving rise to a PCR product of variable size (usually ranging from ~600 to ~700 bp) depending on the repeat number. Finally, the transgene-specific DM-R primer and the insertion site-specific mDmtD-GR primer only generate a DNA sequence if the transgenic mouse belongs to the *Dmt-D* line. A positive DM-R/mDmtD-GR reaction serves therefore to specifically identify

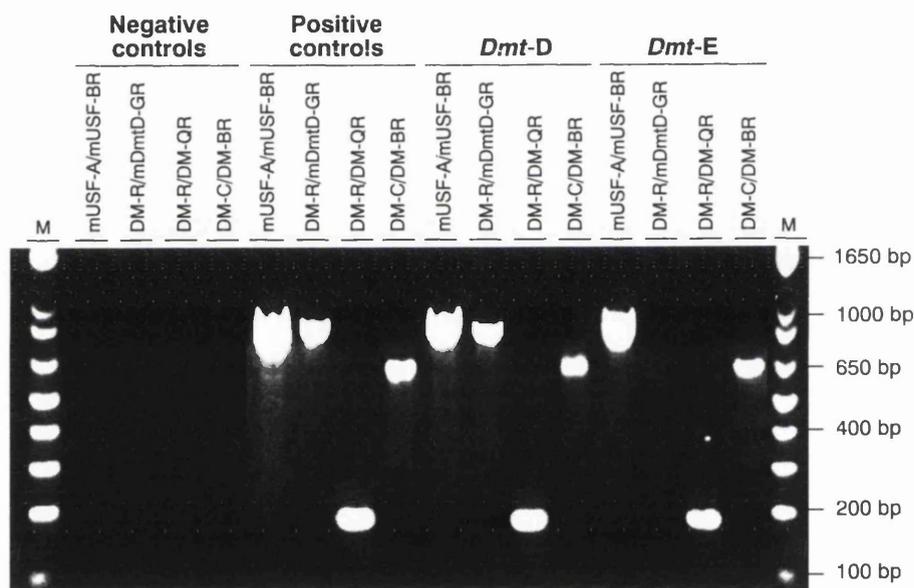
---

<sup>3</sup> Mouse mating pairs were set up by Teresa Fortune (Division of Molecular Genetics, University of Glasgow, Glasgow, UK).

*Dmt-D* transgenic mice. The annealing temperatures for each primer combination, and the respective number of amplification cycles are indicated in Table 3.1.

**Table 3.1. Annealing temperatures and number of cycles for each primer combination used in the *Dmt* mouse PCR genotyping analyses.**

Primer combination	Annealing temperature	Number of cycles
mUSF-A/mUSF-BR	63°C	28
DM-R/DM-QR	68°C	30
DM-C/DM-BR	68°C	30
DM-R/mDmtD-GR	68°C	30



**Figure 3.1. Mouse genotyping for the *Dmt162* transgene.**

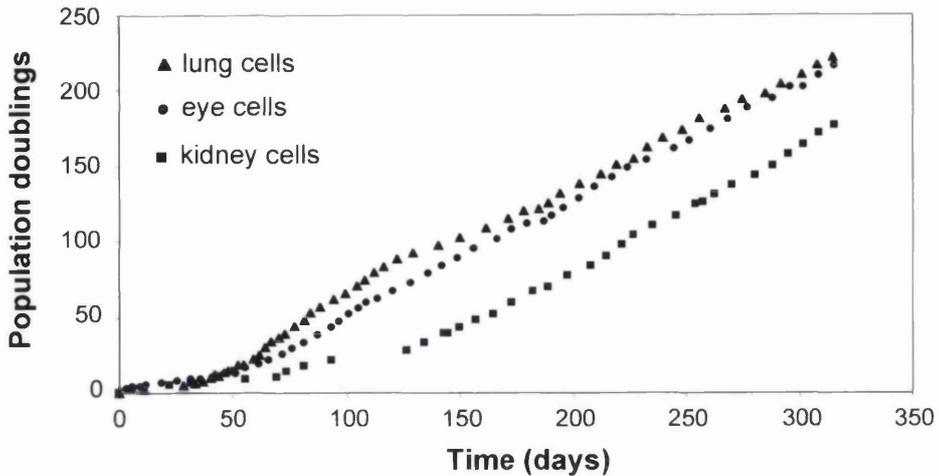
The ethidium bromide-stained agarose gel shows the PCR products obtained from the amplification of three different mouse DNA samples, plus negative control reactions without template DNA. The sets of oligonucleotide primers used in each PCR amplification are shown above each lane. The positive controls consist of a DNA sample extracted from a *Dmt-D* mouse previously genotyped. The two mice analysed in this gel both carried a *Dmt162* transgene in their genome, since a PCR product was generated upon amplification with primers DM-R and DM-QR. The transgenic CAG•CTG repetitive sequence was amplified with primers DM-C and DM-BR. The two mice differed in the insertion site of the *Dmt162* transgene in their genome: one animal was classified as *Dmt-D*, since its tail DNA gave rise to a PCR product when amplified with primers DM-R and mDmtD-GR; whereas the other belonged to the *Dmt-E* line, with no PCR product generated in the same reaction. The sizes of molecular weight markers (M) are shown on the right. PCR analyses were performed as described in Section 2.5.5, using the annealing temperatures listed in Table 3.1. The PCR products were resolved through a 1.5% (w/v) agarose gel, as described in Section 2.5.7.1.

### 3.2.2. Growth dynamics of mouse cell cultures

In an attempt to establish primary mammalian cell cultures in which the dynamics of expanded CAG•CTG repeat tracts could be monitored and investigated *in vitro* over long periods of time, lung, eye and kidney tissue samples were originally harvested from a six-month-old male *Dmt-D* transgenic mouse, since this is the time point at which obvious tissue-specific differences can be easily detected. Using either the explant technique (eye, Section 2.4.1.1) or enzymatic dissociation procedures (lung and kidney, Section 2.4.1.2), mouse cell cultures were successfully established. Cell lines were identified by a designated letter referring to the transgenic mouse line (D, *Dmt-D*; E, *Dmt-E*), followed by a four-digit number that identifies the mouse they were established from, and a final letter corresponding to the progenitor tissue (L, lung; E, eye; K, kidney) (Table 3.2). Following an initial period characterised by low growth rates, all the D2763 cell cultures entered a continuous exponential growth phase, proliferating at similar and constant rates (Figure 3.2). The cell proliferative capacity of each culture was estimated according to the population doubling time (PDT), which was calculated based on the cell counts determined at each passage (Table 3.2). The late rapid cell growth observed *in vitro* is consistent with the spontaneous immortalisation of cells, which is known to occur at a relatively high frequency with mouse cell cultures (Meek *et al.*, 1977; Todaro and Green, 1963). The time taken to reach peak growth rate was longest for the kidney culture, which took around 120 days, in contrast to the lung and eye cell cultures, which took around 60 days.

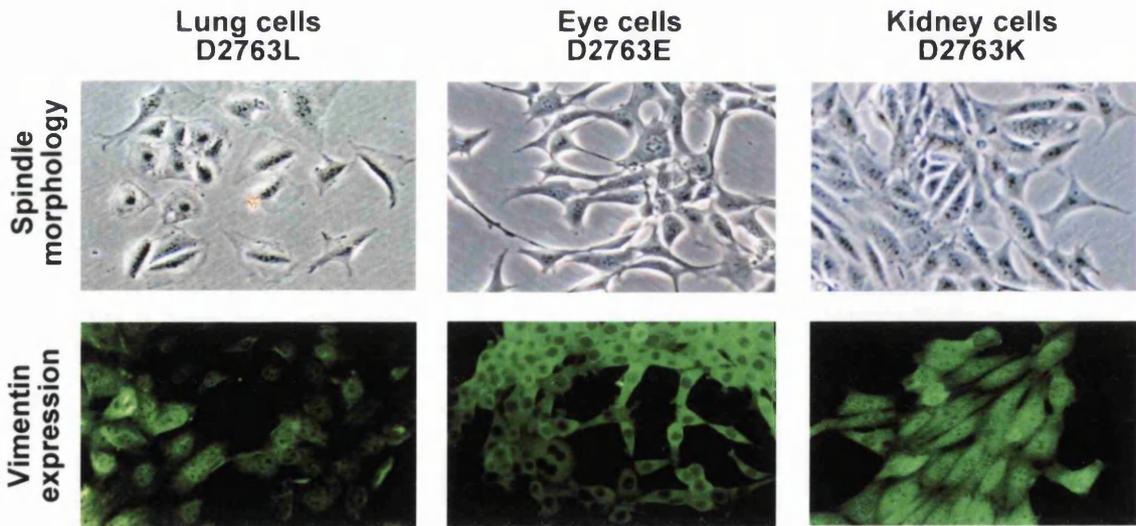
**Table 3.2. Population doubling times of cell lines and relative trinucleotide repeat stability.** The proliferative capacity of each cell culture was estimated based on the population doubling times (PDT), which was calculated as a function of the cumulative cell numbers determined at each passage. Trinucleotide repeat instability was assessed by sensitive SP-PCR techniques.

Mouse genotype	Mouse age	Tissue	Cell line name	PDT	Trinucleotide instability
<i>Dmt-D</i>	6 months	lung	D2763L	32 hours	low
"	"	eye	D2763E	30 hours	medium
"	"	kidney	D2763K	31 hours	high
<i>Dmt-D</i>	5 weeks	kidney	D3111K	42 hours	low
<i>Dmt-D</i>	3 months	kidney	D2967K	45 hours	medium
<i>Dmt-D</i>	30 months	kidney	D979K	29 hours	very high
<i>Dmt-E</i>	8 months	kidney	E3994K	54 hours	very low



**Figure 3.2. Growth dynamics of D2763 cell cultures.**

The graphs show growth curves for the three cell lines established from a six-month-old *Dmt-D* male mouse: triangles (▲), lung cells; circles (●), eye cells; squares (■), kidney cells. For each culture the cells were counted at each passage, the population doubling time was calculated and the cumulative number of population doublings plotted as a function of days in culture.



**Figure 3.3. Fibroblastic phenotype of D2763 cultured cells.**

The pictures show the phenotype of cultured lung, eye and kidney cells after 20 population doublings, derived from a six-month-old *Dmt-D* mouse. The top row shows light micrographs. Note the prevalent fibroblastic spindle morphology. The bottom pictures represent the immunofluorescence detection of vimentin. Cells were stained with a primary mouse monoclonal antibody raised against human vimentin, and a secondary anti-mouse IgM-FITC conjugate. All cells stained positive. All cells stained negative for a panel of cytokeratins.

### **3.2.3. Characterisation of cultured cell types**

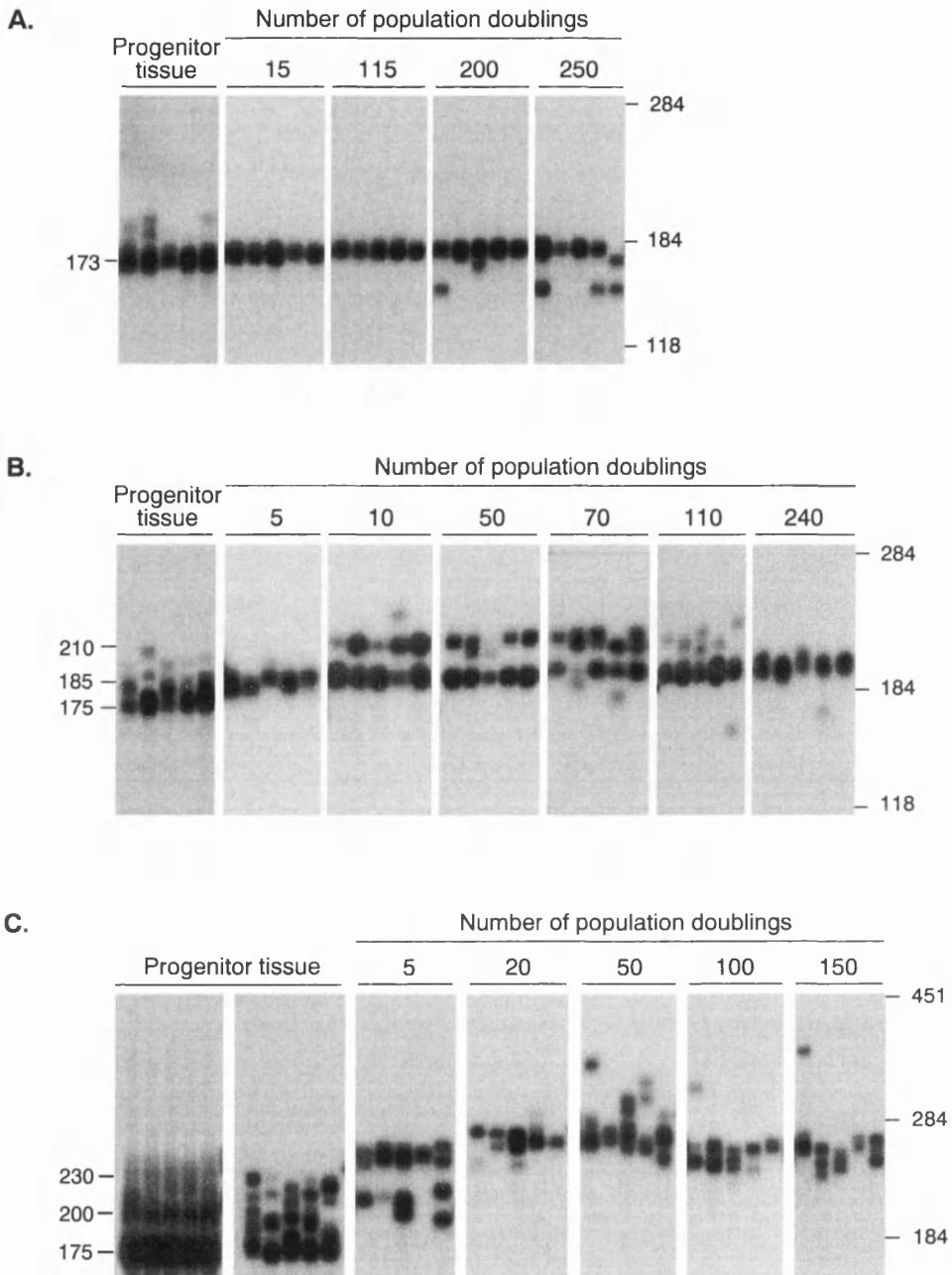
Each culture consisted of an initial heterogeneous cell population, comprising cells with clearly different morphologies and probable distinct proliferative capacities. The cultures soon became more homogeneous, exhibiting prevalent spindle morphology, typical of fibroblasts, after less than five passages (Figure 3.3). The nature of the cultured cells was confirmed immunocytochemically by staining with primary antibodies directed against vimentin and cytokeratins (Herrmann and Aebi, 2000; Schmid *et al.*, 1979; Spector *et al.*, 1998). All the cultures stained positive for vimentin and negative for a panel of cytokeratins, consistent with a fibroblastic, rather than epithelial phenotype (Figure 3.3). Nonetheless, the exact morphology was clearly distinct between the cultures derived from the different tissues. In particular the precise pattern of cytoplasmic staining was not exactly the same for all the cultures investigated. This observation may indicate a different absolute origin for each culture.

### **3.2.4. Tissue-specific trinucleotide instability and selection for longer alleles in cultured mouse cells**

Following the establishment of these tissue-specific *Dmt-D* murine cell lines, DNA samples were collected at every passage and transgene repeat length variability assessed by SP-PCR analysis (Jeffreys *et al.*, 1994; Monckton *et al.*, 1995). The length of the repeat in the cultures and the level of variation were compared with the progenitor allele length in the donor mouse (173 repeats, as determined by PCR analysis of tail DNA at weaning) and the level of variability present in the tissue from which the culture was originally derived.

#### **3.2.4.1. Repeat size variability in lung cell cultures (D2763L cell line)**

The lung tissue from which the culture was established showed relatively low levels of variability with most alleles (>90%) remaining within  $\pm 10$  repeats of the progenitor allele (173 repeats) (Figure 3.4.A). After 15 doublings the lung cell culture displayed an even lower level of variability with the vast majority of cells within  $\pm 5$  repeats of the predominant allele (~175 repeats). The reduction in variability observed in the progression from *in vivo* to *in vitro*, suggests that only a very few cells grew in culture. Indeed, the degree of repeat length homogeneity observed in the culture and the relative increase in size detected, suggest that this culture may have very quickly been taken over by derivatives of a few, or possibly even only one, of the cells present in the original tissue carrying a slightly larger allele. Surprisingly, even after as many as 100 or 200 doublings *in vitro* (corresponding to more than 300 days), the level of repeat variability remained very low with only a small increase in average allele length up to 177 repeats. The maintenance of such a low level of variability *in vitro* after so many doublings indicates that mammalian cells are capable of faithfully replicating large expanded CAG•CTG repeat tracts through numerous cell divisions, even at a locus that is extremely unstable in other cells. Also of note was the relatively late appearance in the culture of a subset of cells carrying a deletion of around 30 repeats relative to



**Figure 3.4. Repeat length variation in *Dmt-D* cultured cells from a six-month-old mouse.** SP-PCR techniques were used to monitor transgene repeat length variation in DNA samples extracted from cultured (A) lung, (B) eye and (C) kidney cells harvested from a six-month-old *Dmt-D* male mouse. Five to ten copies of genomic DNA, collected at different passages were amplified in five independent SP-PCR. Representative SP-PCR amplifications of DNA isolated from the progenitor tissues are shown on the left; each reaction containing around 20 copies of the transgene, except for the first kidney panel, in which around 50-100 molecules of DNA were amplified. The scale on the right displays the molecular weight markers converted into CTG repeat numbers. The repeat lengths of the major PCR products detected are shown on the left.

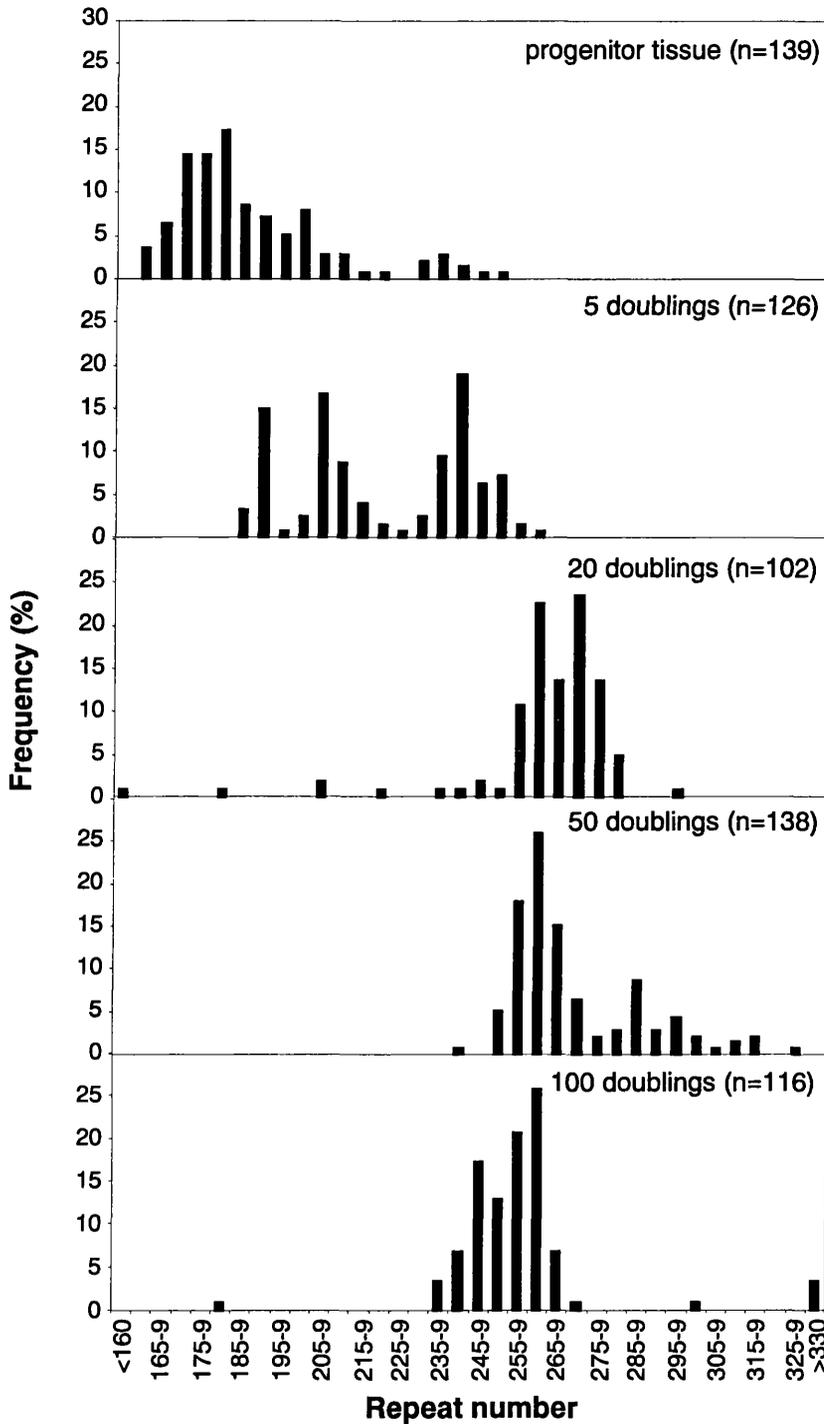
the major allele in the culture. After 200 doublings these cells comprised approximately 10% of the cell population, but rose to ~25% by 250 doublings. Presumably, this shift was mediated by drift and/or selection within the cell population rather than repeated mutations to the same length allele. This assertion is supported by the bidirectional nature and different sizes of similar shifts observed in other cultures.

#### **3.2.4.2. Repeat size variability in eye cell cultures (D2763E cell line)**

Higher levels of instability were detected in cultured eye cells (Figure 3.4.B). Once again, very early in the establishment of the culture, the range of variability observed was very different to that detected in the progenitor tissue. By as few as 5 doublings, the predominant cell population contained an average allele length of ~185 repeats, 12 repeats larger than the major allele in the original tissue (173 repeats). By 10 doublings a distinct subpopulation of cells appeared, with an average size of ~210 repeats. This population presumably represents clonal expansion of a rare cell carrying a large expanded allele, either present in the original tissue or having arisen as a spontaneous mutant *in vitro* during one of the earlier passages. Nonetheless, by 10 doublings each population showed moderate size variability, characterised by small changes mostly limited to fewer than  $\pm 5$  repeats around the average repeat length. The moderate instability was maintained, with the level of variability and the average allele length gradually increasing within each of the two main populations of cells. This effect was very clear up to 70 doublings, by which time the average allele size in each population had risen by a further five repeats relative to that observed at 10 doublings. After 110 doublings however, the proportion of cells in the population carrying the larger alleles started to decrease. By 240 doublings *in vitro*, the eye cell population initially carrying ~185 repeats had increased in average allele length to ~195 repeats, but had overgrown the culture causing a reduction in overall repeat length variability.

#### **3.2.4.3. Repeat size variability in kidney cell cultures (D2763K cell line)**

Kidney is the tissue that shows the highest levels of variability and the largest expansions *in vivo*. Repeat variability within the original tissue showed the typical trimodal distribution highly biased towards expansion (Fortune *et al.*, 2000) (Figure 3.4.C). Most cells carried repeats within the first peak of variability with alleles within -5 to +10 repeats of the progenitor (173 repeats). Additional peaks of variability were observed most clearly at ~200 and ~230 repeats. Although a relatively high level of variability was retained after 5 doublings *in vitro*, the predominant alleles from the first peak of variability detected in the original kidney tissue were entirely absent. It appears as if only cells from the second and third peaks were able to grow *in vitro*. Indeed, by as few as 20 doublings, the second peak of cells had disappeared and only cells carrying the largest repeats were maintained. Single molecule analyses allowed us to quantitatively define more clearly the progression of repeat variability in cultured kidney cells (Figure 3.5). After 20 doublings a narrow range of repeat sizes was defined, with an average repeat length of ~260 repeats, which not only corresponded to the longest repeats initially found *in vivo*, but also overlapped with the largest



**Figure 3.5. Repeat length distributions in cultured kidney cells.**

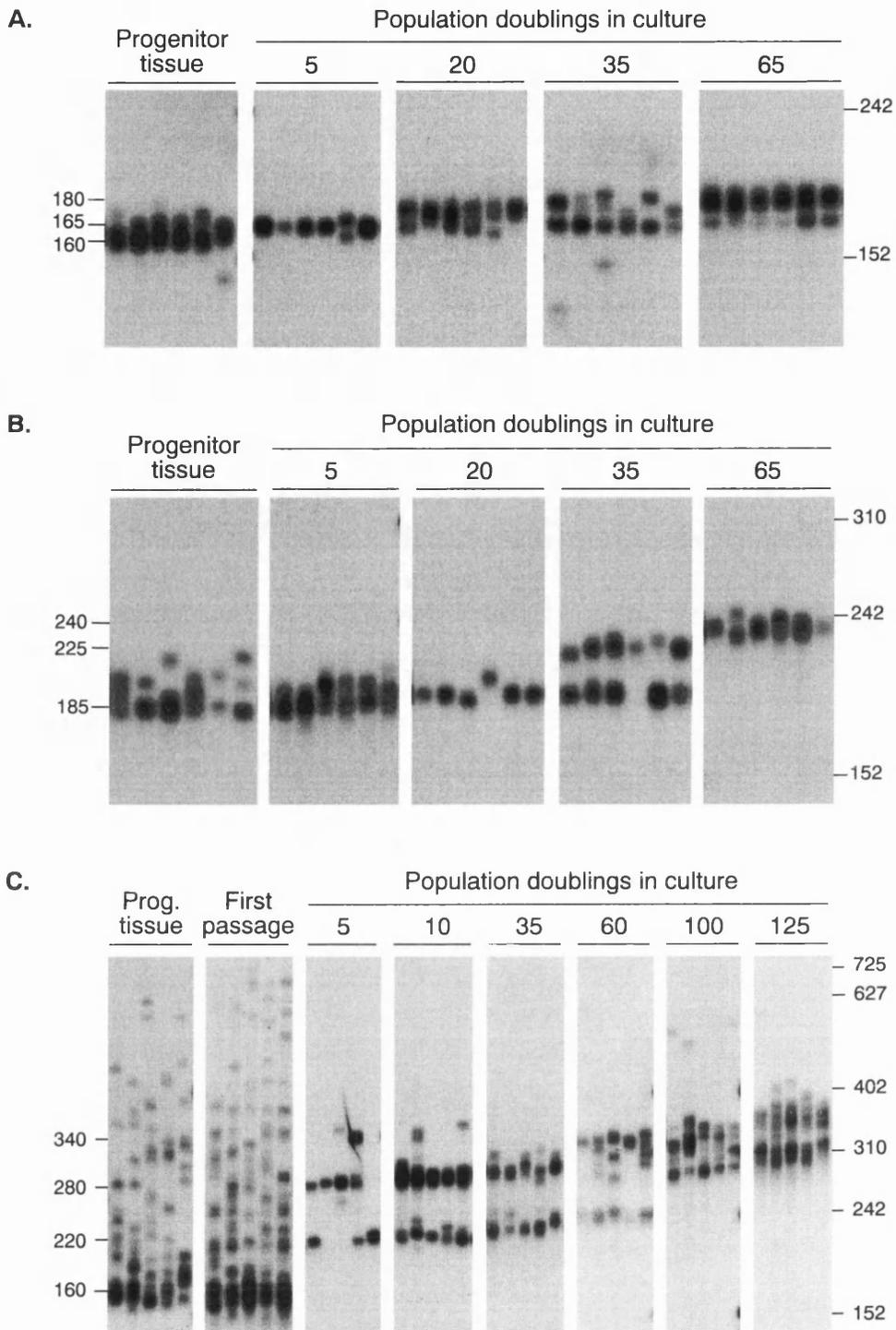
The graphs display the repeat length distributions observed in cultured kidney cells harvested from a six-month-old *Dmt-D* male mouse. For each time point at least 100 individually amplified transgene repeat tracts (n) were sized by SP-PCR at low DNA concentration (1-3 DNA molecules per reaction). Allele lengths were grouped into five repeat size ranges.

peak described above after five doublings in culture. After 50 doublings, a pattern of instability very reminiscent of the highly positively skewed distributions observed *in vivo* in both humans (Monckton *et al.*, 1995) and mice (Fortune *et al.*, 2000) was observed, with very few alleles being detected below the lower boundary of ~250 repeats. The level of repeat variability peaked at this stage, with a mean allele length of 270 repeats and with ~5% of cells containing alleles longer than 300 repeats in length. However, the repeat length heterogeneity was reduced dramatically and the mean allele length decreased to 260 repeats after 100 divisions in culture. A similarly low level of variability was retained even after 150 population doublings. The repeat dynamics in cultured kidney cells suggests the development of an expansion-biased mechanism *in vitro*, which may be disturbed by population fluctuations, such as the dramatic selective sweeps described at 20 and 100 doublings.

It is important to note that no obvious relationship was found between the repeat instability in the three tissue-specific cell lines analysed and their proliferative capacity, as assessed by the population doubling times (Table 3.2). Taken together, the results reported above for the three different tissues, support the maintenance of *in vivo* tissue-specific trinucleotide instability *in vitro*, which is expansion-biased, particularly during the early stages of the eye and kidney cell cultures.

### **3.2.5. Age-of-donor effect on trinucleotide repeat instability observed in culture**

To investigate the influence of the age of the donor mouse at sacrifice on the trinucleotide instability detected in culture, kidney cell lines were established from two younger *Dmt-D* male mice, aged five weeks (D3111K cell line) and three months (D2967K cell line), and one very old mouse aged 30 months (D979K cell line). The two cultures from the young mice showed similar growth rates, with population doubling times of 42 and 45 hours, respectively (Table 3.2). Consistent with the age-dependent accumulation of somatic mosaicism *in vivo*, the youngest mouse analysed in this study displayed very little somatic variation in the original tissue from which the cell line was derived, with most cells with an allele size of ~160 to ~175 CTG repeats (Figure 3.6.A). As with previous cultures though, the level of variability in the D3111K cell culture was almost immediately reduced in the first few passages and by 5 doublings the average allele size was ~165 repeats with most variants within  $\pm 5$  repeats. The range of variation and average allele length increased up to 35 doublings, by which point two major populations of cells with ~170 and ~185 repeats were present. After an additional 30 doublings the cells containing the larger alleles predominated, accompanied by an overall reduction in the range of variability, reflective of a selective sweep. The repeat sizes ranged from ~185 to ~215 repeats in the original kidney tissue of the three-month old mouse, intermediate between that observed in the five-week and six-month old mice (Figure 3.6.B). Once again, variability was rapidly reduced in the D2967K culture: most notably by 20 doublings the average allele length had increased to ~195 repeats, but with a reduced overall range. By 35 doublings a second major population of cells carrying approximately 225 repeats appeared. This population took over the culture completely by 65 doublings and increased



**Figure 3.6. Age-of-donor effects on repeat length variation in cultured kidney cells.**

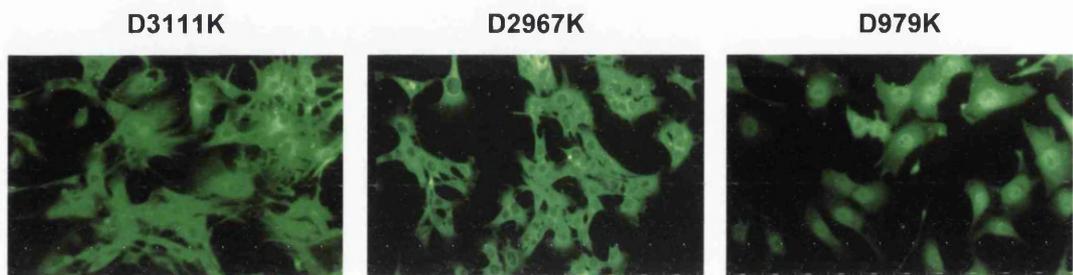
The autoradiographs shown are representative SP-PCR analyses of DNA samples collected from cultured kidney cells harvested from (A) a five-week-old *Dmt-D* mouse, D3111K cell line; (B) a three-month-old *Dmt-D* mouse, D2967K cell line; and (C) a 30-month-old *Dmt-D* mouse, D979K cell line (see Table 3.2). Representative SP-PCR amplifications of DNA isolated from the progenitor kidneys are also presented on the left, each reaction containing around five to 50 copies of the transgene. The molecular weight markers are shown on the right, after conversion into number of CTG repeats. The length sizes of the main PCR products detected are shown on the left.

in length up to ~240 repeats.

The level of variation present *in vivo* in the very old mouse was, as expected, very high with a small subset of cells carrying alleles as large as 720 repeats (Figure 3.6.C). Nonetheless, the most predominant cells contained expansions in the range of 160 to 180 repeats. This high level of variability though, was massively reduced very early in the establishment of the D979K culture. By five and 10 doublings only three major, but highly distinct, populations persisted suggesting that this culture was most likely derived from only three progenitor cells, carrying transgenic sequences with average sizes of 220, 280 and 340 CTG repeats. Once again the allele sizes present in the cells that predominated in culture were much larger than present in the majority of cells *in vivo*. As the culture was continuously passaged, not only the repeat size variability within each subpopulation but also the mean repeat length appeared to increase significantly. However, the population of cells carrying 220 repeats disappeared after 100 doublings, following a rapid expansion of the transgene by 20 units (up to 240 CTG repeats) over as few as ~60 population doublings in culture. After 100 doublings the upper population of cells displayed major repeat size heterogeneity, with some alleles containing up to 600 repeats. These repeats appear to undergo rapid expansion-biased mutation, resulting in a dramatic increase in the mean repeat number over as few as 25 population doublings *in vitro*: from ~315 repeats after 100 doublings, up to ~355 repeats after 125 doublings. Interestingly, no evident selective sweeps were observed over the last two time points analysed for this culture, in contrast to the results obtained with the D2763K cell line, and indeed supporting a high propensity for rapid expansion of the CAG•CTG repetitive tract in this particular cell line (D979K).

The difference in repeat size variation observed between the cell lines established from three different aged mice could not obviously be accounted for by the nature of the cell types grown from each animal. All cells cultured *in vitro* displayed a characteristic spindle morphology, and expressed vimentin filaments, as detected by immunocytochemistry techniques (Figure 3.7), in agreement with their fibroblastic phenotype. Nevertheless, as observed for D2763 cell lines, careful examination revealed subtle differences regarding the exact morphology of the cells and patterns of cellular staining, particularly between D979K cells and the other two cell lines. Whether or not this observation is relevant, and indicates genuine differences in the absolute origin of each culture remains unclear.

Overall, these data confirm the proliferative advantage *in vitro* of cells that contained large repeats *in vivo*. They also indicate that the expansion-biased progression of somatic instability *in vitro* is a reproducible phenomenon using kidneys from *Dmt-D* mice as a source material and that selective sweeps are a common occurrence. These data also appear to support an effect of the age of the mouse at sacrifice on the stability of the transgene in culture, with the repeat appearing to be less stable in cell lines derived from older mice than those from younger mice. However, this may be a result of the longer allele lengths that predominate in the cultures derived from older mice, rather than a true age-of-donor effect on repeat stability.



**Figure 3.7. Fibroblastic phenotype of cultured kidney cells from different aged mice.**

The pictures represent the immunostaining of kidney cells derived from three different mice aged three weeks (D3111K cells), six months (D2967K cells) and 30 months (D979K cells). Vimentin filaments were detected with a primary mouse monoclonal antibody raised against human vimentin, and a secondary anti-mouse IgM-FITC conjugate. All cells stained positive for vimentin, but negative for a panel of cytokeratins.

### **3.2.6. Accumulation of mutations in single cell-derived clones from kidney cell cultures**

To discount the possibility that the variability detected in culture could be an artefact derived purely from an *in vitro* selection process for the cells harbouring longer alleles, rather than the outcome of the intrinsic instability of the transgene *in vitro*, clonal lines derived from single cells were established. The clones were isolated by limiting dilution from kidney cell cultures established from two *Dmt-D* mice aged six and 30 months: D2763K and D979K, respectively. D2763K cells had been growing for 20 population doublings, whereas D979K had been growing for only 10 doublings prior to the cloning step. All the clones were expanded for a further 20 doublings and the CTG repeat size variability was subsequently assessed by SP-PCR. In this experiment, if the progenitor cell, which was clonally expanded, carried a stable allele, low levels of repeat size variability would be detected in the final culture. On the other hand, if the original cell carried an unstable repeat tract, a high degree of trinucleotide repeat length heterogeneity would be expected at later stages. If, however, the culture resulted from the expansion of more than one cell, we would expect to see distinct subpopulations of variants. Therefore, clones that showed two or more clear independent subpopulations of cells carrying different sized repeats were considered to represent either the presence of multiple input cells at the time of cloning or, possibly, new mutant alleles that arose soon after plating. In either case those clones were excluded from the study.

All the clones derived from a six-month-old mouse cell line exhibited repeat size heterogeneity, which corroborates the progression of somatic instability in culture over as few as 20 population doublings (Figure 3.8.A). In addition, they showed different average repeat sizes and different ranges of repeat distribution. For one clone in particular (clone D2763Kc2), a very broad range of repeat sizes was observed. Such a difference could not be accounted for by cell division rates, since no significant variations in the proliferative capacities were observed between the three clones studied here: all the clones were proliferating at an average rate of one population doubling every 30 hours.

The original D979K kidney cell culture, from which several clones were derived, exhibited three main subpopulations of cells, carrying repeats with average sizes of ~230, ~290 and ~340 CTG units. The repeat sizes of the alleles carried by the ten D979K clones fall within these three major peaks of length variability (Figure 3.8.B), suggesting that each clone was derived from individual cells originally included in those three main subpopulations. One particular clone (clone D979Kc1) carries a transgene sequence longer than any of the alleles previously detected in the progenitor culture under the conditions employed in the SP-PCR analysis (DNA concentration and number of replicates). There is, therefore, the possibility that clone D979Kc1 shows greater propensity for trinucleotide repeat expansion than all the others, leading to a rapid increase in its mean repeat size. Clones D979Kc4 and D979Kc5 both exhibit higher levels of repeat size heterogeneity compared to the remaining clonal cell lines. The possibility exists that these two clones did not expand from a single cell present in the progenitor culture, but from multiple cells



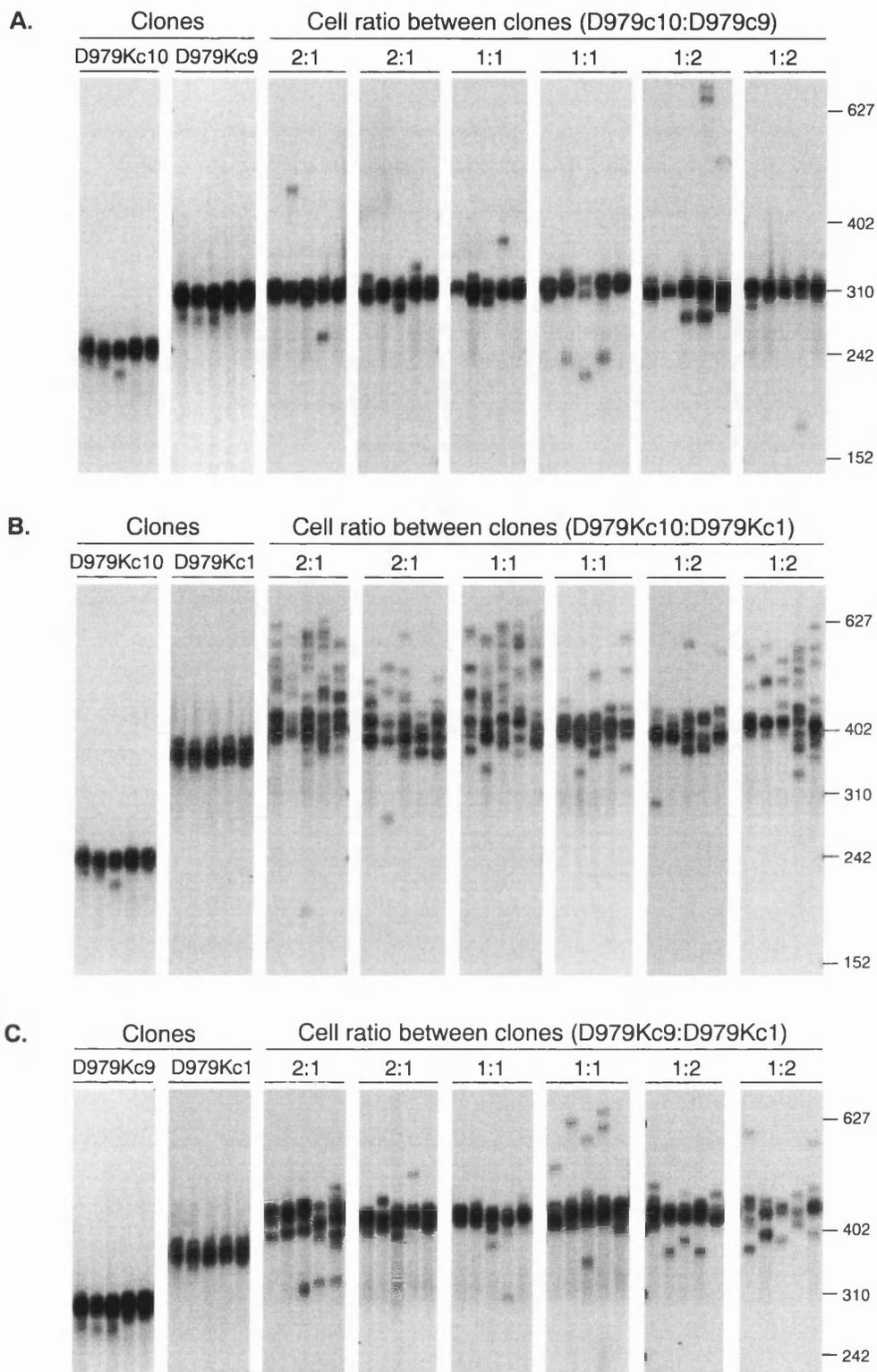
carrying different repeat lengths. As observed for the D2763K clones, differences in the repeat dynamics between D979K clones could not be explained by clonal-specific proliferative capacities, since no significant variations in the population doubling times (~30 hours) were observed between the ten clones studied here.

Interestingly, the repeat size variability within single-cell derived cultures was greater in D2763K clones than in D979K clones (Figure 3.8). This difference may be the result of a greater repeat length heterogeneity previously detected in the progenitor D2763K culture. In contrast, although the parental D979K culture consisted of three distinct subpopulations of cells carrying different sized alleles, little repeat number variation was detected around the average repeat size carried by each subpopulation. Different degrees of somatic mosaicism between these two independent cell lines may be explained by differences in the progenitor cell types, despite the fibroblastic phenotype developed by cultured cells at later stages. Alternatively, the accumulation of distinct mutations over time, particularly during spontaneous immortalisation, most likely to involve genes associated with DNA repair and/or cell cycle surveillance, may also contribute to the different degrees of repeat size variability observed between the two cell lines.

### **3.2.7. Preferential accumulation of longer alleles in competition assays between clones carrying different sized repeats**

Given the preferential accumulation of cells carrying longer alleles, particularly over the first few passages in culture (Figures 3.4 and 3.6), it was our hypothesis that this phenomenon could be mediated by the growth advantage of cells carrying large CAG•CTG repeats. Clonal cell lines were therefore deliberately mixed prior to serial passage and repeat size monitoring by SP-PCR. Competition growth assays were performed with clonal cell lines derived from a 30-month-old *Dmt-D* mouse (D979K clones), carrying different sizes of the CAG•CTG repeat but sharing the same genomic background. Three D979K clones were selected according, not only to their different repeat numbers, but also to their low transgene size heterogeneity, to ensure that clonal expansion had occurred from a single progenitor cell. Clones D979Kc1, carrying ~380 repeats, D979Kc9, carrying ~300 repeats and D979Kc10, carrying ~240 repeats (Figure 3.8.B), were chosen to carry out this study. These clones were mixed together in pairs, at three different cell number ratios, either an equal cell number from each clone (1:1 ratio) or a two-fold excess of one of the clones relative to the other (1:2 and 2:1 ratios). Two replicates of each mixed culture were grown for 40 population doublings *in vitro*, prior to analysis of repeat size variability by SP-PCR techniques (Figure 3.9).

SP-PCR analyses revealed that cells carrying longer alleles overgrew the culture over the period of time monitored in this study and at all the cell ratios tested (Figure 3.9). Regardless of the exact repeat tract length in each clone and the extent of the difference in repeat size between the two clones mixed together (differences varying from 60 to 140 CTG repeats), the cells carrying larger transgene molecules overgrew the culture, even when they were outnumbered by a two-fold excess of cells harbouring shorter trinucleotide repeats. Consequently, following 40 doublings,



**Figure 3.9. Competition growth assays between single cell-derived clones.**

Three kidney cell clones carrying different sized CAG•CTG repeats were mixed in duplicate at three different ratios. The ratios between the numbers of cells from each clone in a mixed culture are shown above each autoradiograph. The autoradiographs show representative SP-PCR products obtained from the amplification of DNA samples extracted from the mixed cultures following 40 population doublings in culture. An average of 5-20 DNA molecules were amplified in each reaction. The scale on the right displays the molecular weight markers converted into CTG repeat numbers.

very few short alleles were detected in culture. Particularly interesting, yet intriguing, are the results obtained with clone D979Kc1 (~380 repeats), which exhibited clear preferential growth survival in culture, and consistent higher levels of expansion-biased repeat instability when mixed together with clone D979Kc10. D979Kc1 cells showed evidence of rapidly expanding unstable repeat tracts, as illustrated by the rapid increase of ~30 units in repeat length over a period of 40 doublings in culture.

The probability that one clone would completely outgrow another in six independent replicate cultures, regardless of their relative repeat sizes (two-tailed analysis), would be  $1/2^5$  (1/32). If the relative repeat sizes are taken into account, the probability that the clone carrying longer alleles would overgrow another clone carrying shorter repeats (one-tailed analysis) would then be  $1/2^6$  (1/64). Both numbers are low enough to consider that the relative growth survival for the three clones tested was unlikely to have occurred by chance. It may therefore be concluded that the selective growth advantage follows the order: D979Kc1 > D979Kc9 > D979Kc10, which happens to parallel the average repeat size rank. Nevertheless, given the low number of clones tested in this experiment, the probability of the correlation between preferential growth and allele size to occur by chance ( $1/3! = 1/6$ ) is not low enough to make this observation highly significant.

In summary, competition assays between three independent single cell-derived clones established from the same progenitor culture, which proliferated at similar rates, suggested a parallel between the average repeat length and preferential growth survival in culture. However, the analysis must be extended to additional clones to increase the significance of this result, and provide further evidence in support of a growth advantage for cells harbouring longer CAG•CTG repeat tracts in culture.

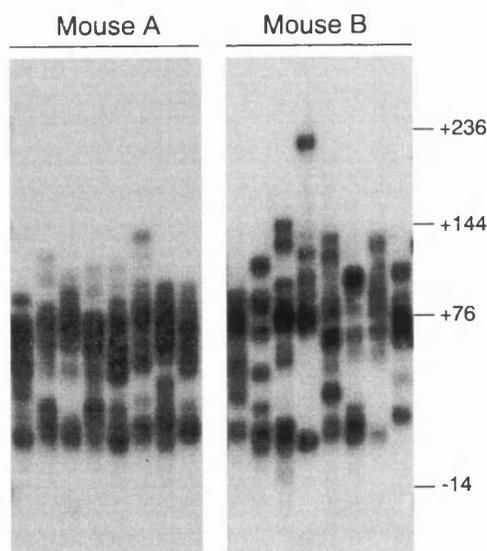
### **3.2.8. Repeat size distributions in kidney and liver tissue samples in vivo**

Three cell lines, established from independent organ samples collected from one single six-month-old *Dmt-D* mouse, revealed different repeat dynamics *in vitro*, suggesting the replication of the tissue-specific character of somatic mosaicism in culture. Differences in cell line-specific trinucleotide repeat instability *in vitro* could not be accounted for by the observed nature of the cells grown from each tissue. Immunocytochemical characterisation of the cultured mouse cells has revealed a typical fibroblastic phenotype, characterised by vimentin expression and lack of cytokeratin staining. However, one cannot reject the hypothesis of a cell-specific trinucleotide repeat instability mechanism. In fact it has now been established that most cells grown *in vitro* develop a fibroblastic phenotype, including the expression of vimentin- and fibroblast-specific proteins (Pollack *et al.*, 1997), independently of the original cell type they may have derived from. It might be possible that by the time cell characterisation was performed in this study, the cultured cells had already adopted a fibroblastic phenotype. Alternatively, fibroblasts derived from distinct tissues may be sufficiently different to explain the diversity in the trinucleotide repeat stability observed among lung, eye and kidney fibroblastic cell lines. In either case, differences in

trinucleotide somatic repeat instability would be dependent on cell-specific factors. To support this hypothesis, three peaks of repeat size variability have been observed in the kidney of *Dmt-D* mice (Figure 3.4.C) (Fortune *et al.*, 2000). Furthermore, different mouse models that show trinucleotide somatic mosaicism, have displayed a clear bimodal repeat distribution in liver DNA samples (Lia *et al.*, 1998; Mangiarini *et al.*, 1997; Manley *et al.*, 1999a). This pattern of instability in the liver of *Dmt-D* mice has been confirmed by SP-PCR analysis of DNA samples collected at 24 months of age (Figure 3.10).

High input DNA SP-PCR amplifications (10-50 template molecules per reaction) revealed the bimodal distribution of repeat sizes in both liver samples. The first population of cells, consisting of ~20% of the cells, carried short repeats that had only expanded an average of 0-10 units from the progenitor allele. A second population of alleles, with a similar size, was detected, and shown to carry repeats  $70 \pm 5$  repeats longer than the progenitor allele. The two mice analysed here inherited progenitor alleles that differed 20 repeats in size, which may explain the apparent higher levels of expansion-biased somatic instability observed in mouse B (progenitor allele: ~188 repeats) compared to mouse A (progenitor allele: ~166 repeats)

In summary, *Dmt-D* mice, in agreement with other murine models of trinucleotide repeat instability found in the published literature, exhibit a characteristic bimodal pattern of repeat size variability in the liver, suggesting a possible cell type-specific somatic instability and a role for cell-specific factors controlling somatic mosaicism, as previously proposed for kidney (Fortune *et al.*, 2000).



**Figure 3.10. Trinucleotide repeat length variability in liver tissue samples from two 24-month-old *Dmt-D* mice.**

The stability of the CAG•CTG transgenic repeat in the liver of two *Dmt-D* mice was assessed by SP-PCR analysis at 24 months of age. The autoradiographs show eight independent SP-PCR amplifications of 10-50 transgene molecules for each sample. The repeat size lengths show a clear bimodal distributions in both mice, with two main peaks of repeat variability. The apparent higher levels of expansion-biased repeat instability observed in mouse B may result from the longer progenitor allele inherited by this mouse (~185 repeats), when compared to mouse A (~165 repeats). The size of the molecular weight markers are shown on the right, after conversion into repeat size differences from the progenitor allele.

### 3.2.9. Trinucleotide repeat stability in *Dmt-E* kidney cells

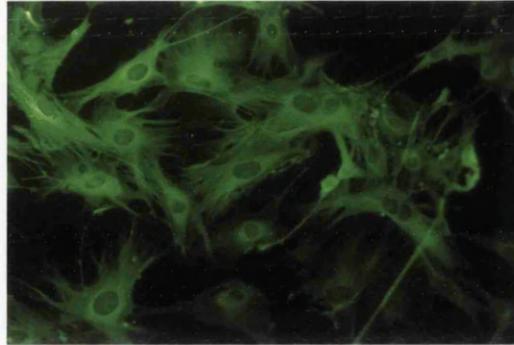
The instability of the *Dmt162* transgene is highly dependent on positional factors, given the major effect of the integration site on the levels of somatic mosaicism detected in the different mouse lines. While the CAG•CTG transgenic repeat shows great levels of somatic mosaicism in *Dmt-D* mice, it appears to remain remarkably stable in the somatic tissues of *Dmt-E* animals (Fortune *et al.*, 2000). To test if trinucleotide repeat dynamics would be similarly controlled by *cis* factors in an *in vitro* environment, particularly, if low levels of repeat instability would still be observed in *Dmt-E* cell cultures, a kidney cell line was established from an eight-month-old *Dmt-E* male mouse by enzymatic dissociation as described before (Section 2.4.1.2). This culture consisted of cells with a fibroblastic phenotype, as determined by immunostaining of vimentin intermediate filaments (Figure 3.11).

SP-PCR analyses were carried out to monitor the repeat size dynamics over time in this cell line (Figure 3.12). As expected, the original tissue exhibited extremely low levels of somatic trinucleotide instability, with all the cells carrying around 156 CAG•CTG repeats. A total lack of repeat length mutations was apparent throughout the culture, with no major expansions or contractions being detected at any time point included in this study. All cells appeared to carry the same repeat number, even after 45 population doublings in culture.

The *cis* factors controlling the repeat instability in the somatic cells of the different *Dmt* mouse lines are still retained in culture, giving rise to marked differences in the levels of repeat length variability between *Dmt-D* and *Dmt-E* kidney cell lines. This observation rejects the hypothesis that the instability observed in cultured *Dmt-D* cells might arise from major changes in the DNA metabolism during the transition from an *in vivo* environment to *in vitro* conditions, leading to an overall genomic destabilisation.

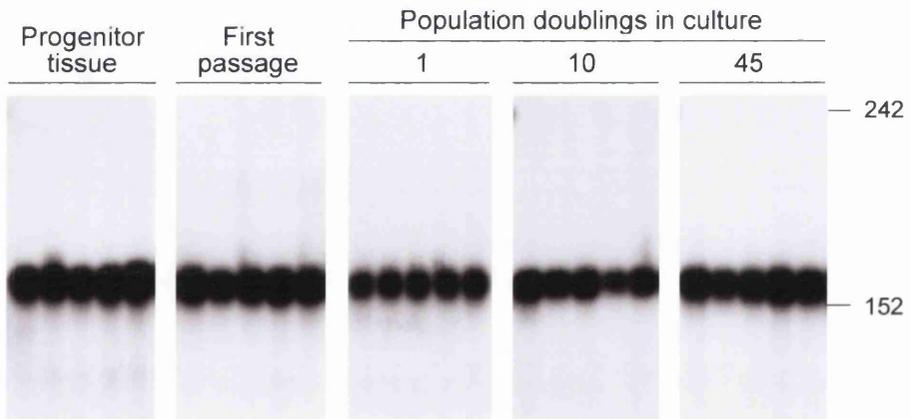
## 3.3. Discussion

The expansion of simple repetitive trinucleotide sequences is causally involved in the molecular bases of an increasing number of human diseases. In the soma, repeat instability is expansion-biased, tissue-specific and age-dependent; dynamics that are consistent with a role in the tissue specificity and progressive nature of the symptoms. However, little is known about the mutation processes that give rise to somatic mosaicism. To facilitate the study of somatic repeat instability in a mammalian system, mouse models of unstable CAG•CTG trinucleotides were previously created (Monckton *et al.*, 1997), one line of which (*Dmt-D*) replicates gross age-dependent, tissue-specific, expansion-biased somatic mosaicism (Fortune *et al.*, 2000). Tissue samples have now been harvested from these mice to establish cell lines, and monitor repeat stability over a long period of time under conditions of rapid cell proliferation. These investigations have revealed that the dynamic pathway observed *in vivo*, *i.e.* accumulation of multiple small mutations biased towards expansions, is conserved in an *in vitro* mammalian cell model. Most



**Figure 3.11. Fibroblastic phenotype of cultured kidney cells from a *Dmt-E* mouse.**

The pictures show the presence of vimentin intermediate filaments in a kidney cell culture established from a *Dmt-E* mouse. A primary mouse monoclonal antibody raised against human vimentin, and a secondary anti-mouse IgM-FITC conjugate were used in the immunostaining. All cells stained positive for vimentin. No cytokeratin expression was observed in the same cells.



**Figure 3.12. Trinucleotide repeat stability in cultured kidney cells from a *Dmt-E* mouse.**

The repeat dynamics in a kidney cell line established from an eight-month-old male *Dmt-E* mouse was monitored by SP-PCR analysis over time. The autoradiographs show representative SP-PCR products obtained from the amplification of 10-20 copies of genomic DNA, collected at different passages. Five independent amplification reactions are shown for each time point. The scale on the right displays the molecular weight markers converted into triplet repeat numbers.

interestingly, the tissue specificity observed *in vitro* also reflects that observed *in vivo*. In particular, the repeat remains remarkably stable in lung cells, even after several hundred days and hundreds of cell divisions in culture. In contrast, the repeat is very unstable in cultured kidney cells. These tissue-specific differences are observed, despite the fact that the population doubling time for cell lines from the three different tissues is similar. It has been postulated that mitotic cycles generate somatic heterogeneity of repeat lengths. According to this model, it would be predicted that the higher the cellular proliferation rates, the higher the variability in repeat number generated. Although the tissue specificity of somatic mosaicism *in vivo* shows no obvious correlation with the known rates of cell turnover, it could be argued that the multi-cell type complexity of whole tissues might have masked such a relationship. The results presented from an *in vitro* model demonstrate conclusively that variation in repeat stability cannot be accounted for simply in terms of population doubling times and must be related to other cell type-specific effects. Cell lines homogeneous for cell type, such as those described here, should greatly facilitate the identification of such factors.

Evidence for cell-specific dynamics was revealed by the detailed analysis of repeat variability in the liver of old *Dmt-D* mice. The repeat length distributions showed two distinct cell populations with differing repeat dynamics. Interestingly, a characteristic bimodal pattern of mutation has also been observed in liver samples from R6/1 mice transgenic for exon 1 of the human *HD* gene (Mangiarini *et al.*, 1997; Manley *et al.*, 1999a). This observation may serve as evidence for a role for cell-specific, as well as tissue-specific factors in the dynamics of trinucleotide repeats. The liver shares nervous, lymph, vascular and fat-storing systems with tissues in which the repeat remains relatively stable. It is therefore reasonable to speculate that these cell types are not the origin of the great liver-specific expansions. Instead, it may be speculated that the shared cell types carry stable repetitive tracts. Given the characteristic two peaks of trinucleotide instability detected in liver samples, at least one highly specialised liver cell type might exhibit great expansion-biased repeat instability and account for the accumulation of longer transgenic alleles. Two cell lineages that are unique to this tissue, such as the liver-specific epithelial cell (the hepatocyte) and the Kupffer cell (liver-specific phagocytic cell), can be considered as good candidates. Both cells types exhibit high metabolic rates, and possibly high levels of oxidative stress (Kono *et al.*, 2000; Wheeler *et al.*, 2001), which may be associated with high levels of triplet repeat instability (Chapter 5), and with the development of the characteristic mutation profile in liver samples. Similarly trimodal repeat distributions have been reported in the kidney of *Dmt-D* mice, and likewise it was speculated that nephron cell types, present specifically in that organ, would be the underlying cause of this observation (Fortune *et al.*, 2000). Microdissection and/or cell sorting procedures should be used to address this issue in further detail. Indeed, single cell PCR analyses, associated to laser microdissection techniques, were carried out on human DRPLA brains and revealed cell-to-cell differences in the CAG•CTG repeat number. Again, the distinct patterns of somatic mosaicism detected between neuron types could not be explained by the differences in the number of cell division cycles alone, and the authors claimed a major role for cell-specific modifiers of somatic mosaicism (Hashida *et al.*, 2001).

Surprisingly, cells carrying longer repeats were selected during the first few passages of all of the cultures, quickly giving rise to a larger average size than observed in the progenitor tissue. This increase was accompanied by an overall reduction in variability, suggesting that each cell line was expanded from a very small number of cells. The probable explanation for this effect is the selective growth advantage in culture of cells that already contained a longer repeat *in vivo*. Alternatively, it is possible, that the highly proliferative cells from which the cultures were established had a very high rate of expansion during their establishment. However, the fact that this effect was observed in all of the cultures, but was most prominent in cultures established from tissues with a high degree of mosaicism already established *in vivo*, suggests the former possibility is most likely. Similarly, EBV-transformed lymphoblastoid cell lines harbour larger repeats than peripheral blood leukocytes (Ashizawa *et al.*, 1996), also supporting a growth advantage *in vitro* of cells that contained longer repeat tracts *in vivo*. The factors that might co-facilitate repeat expansion *in vivo* and rapid proliferation *in vitro* remain unknown.

Selective sweeps are not restricted to the early stage of the culture, and continue to occur at later passages. These later sweeps can similarly result in dramatic shifts in repeat length distributions. For instance, the shift from a mean allele size of ~195 repeats after 20 doublings in the kidney cell culture from the three-month-old mouse, to ~240 repeats after 65 doublings does not appear to result from the gradual accumulation of multiple mutations in a homogeneously evolving population (Figure 3.6.B). Rather, it appears that at some point in the culture a rare cell arose with a markedly larger allele than the average, and that derivatives of this cell had a selective advantage, eventually taking over the culture completely. Other less dramatic selective sweeps have been observed at late passages, but they were not always in favour of the cells carrying larger alleles. For instance, the bimodal distribution present in the eye cell culture after 70 doublings (peaks at 190 and 210 repeats), resolved to a unimodal population of ~200 repeats by 240 doublings (Figure 3.4.B). Presumably, the favourable advantages gained by the selected cells resulted from the acquisition of additional mutations elsewhere in the genome, rather than from any direct effect of expanded CAG•CTG repeats on cell proliferation. In contrast to the bidirectional selective sweeps observed in *Dmt-D* cultures, competition growth assays between three *Dmt-D* single cell-derived clones with the same clonal origin, and hence the same genetic background, has consistently revealed a growth advantage for cells carrying longer alleles over a maximum of 40 population doublings, which could not be accounted for by differences in the rates of cell turnover between clonal cell lines. Despite the low number of clones tested, and lack of statistical significance, these results may suggest a parallel between the selective growth, proliferative capacity and the repeat tract length.

Similar selective sweeps have been observed in human EBV-transformed LBCLs carrying GAA•TTC expansions at the *FRDA* locus (Bidichandani *et al.*, 1999), and CTG•CAG expansions at the *DM1* locus (Khajavi *et al.*, 2001). Whilst the selective sweeps observed at the GAA•TTC expansion in the *FRDA* cells were not obviously biased in their direction, those at the *DM1* locus were associated with a growth advantage of cells carrying longer repeats. Indeed, the authors have suggested that there is a direct cause and effect relationship between the CTG•CAG expansion

length at the *DM1* locus and the proliferative capacity of the cell, presumably mediated by abnormal function of one of the genes flanking the endogenous repeat. At present we cannot exclude a positive correlation between the size of the CAG•CTG repeat expansion and *Dmt-D* cell proliferation capacity, as previously reported for DM1 LBCLs (Khajavi *et al.*, 2001). The use of sensitive flow-cytometric techniques of pulse-labelled cells with 5'-bromo-3-deoxyuridine would allow a more precise calculation of the population doubling time, and might reveal subtle differences that might have not been detected with the techniques used in this work. Nevertheless, if confirmed, a correlation between CAG•CTG repeat size and cell turnover in *Dmt-D* cell lines would not be expected to be mediated by one of the genes flanking the endogenous repeat, since the transgene contains no coding sequences, it has been randomly integrated into the mouse genome and it is not associated with any obvious phenotype in either hemi- or homozygous *Dmt-D* mice (see Chapter 9).

More sensitive techniques would also clarify the relationship between the number of population doublings and the number of cell generations. In a cell culture undergoing multiple passages, cells are subjected to a periodic reduction of the population size at each passage and to genetic drift, which increases the likelihood of a selective sweep to occur. Since non-dividing cells accumulate in the culture, dividing cells have to compensate for non-dividers in order to accomplish additional population doublings, thus, individual proliferating cells undergo more divisions, called cell generations, than the number of population doublings. In fact, the number of cell generations can be twice as high as the number of population doublings (Rubelj *et al.*, 1999). The differential proliferation rates within a cell culture may well be on the basis of the selective sweeps observed. Such selective sweeps complicate attempts to provide an accurate measure of the repeat dynamics in any given cell line since the level of variation present can be rapidly reduced by selection. Moreover, the mutations that are selected will be those affecting cell growth and turnover and these might also be expected to indirectly or directly affect DNA metabolism. As such, it is probable that the dynamics of repeat metabolism will change with time. Indeed, it appears as if the repeats are most unstable in the early passages of the culture and become more stable with time. Thus, it appears as if the early passage cells are more *in vivo*-like, initially retaining the cell-type specificity of instability, but progressing toward a similar level of high stability. This further emphasises the requirement for a readily accessible source of tissue as afforded by the transgenic mice. Patient-derived human samples from individuals with rare conditions are not readily available. Moreover, the common genetic background of inbred transgenic mouse lines and their controlled environment reduces variability due to other factors that are not easily corrected when using human samples.

In addition to complicating attempts to quantify repeat dynamics in cells lines, the selective sweeps observed may have other more general implications. Sweeps have been detected given that very high levels of variation are a feature of the system described. There is no reason to assume that such sweeps are not also a feature of other tissue culture systems, and it is important that their potential effects are also considered when attempting to interpret the results obtained. This is particularly critical for attempts to study other aspects of triplet repeat biology such as effects on

gene expression. It is standard practice to measure repeat length in either the progenitor tissue or at one time point in the culture and assume this remains constant. Our results indicate that such simple assumptions cannot be made.

Manley *et al.*, have reported the establishment of a similar mouse tissue culture system using tissues from the R6/2 mice transgenic for exon 1 of the *HD* gene carrying 154-155 CAG repeats. Rather than using primary cell lines though, they have established continuously passaged simian virus 40 (SV40)-transformed fibroblasts (Manley *et al.*, 1999a). This system also provides evidence for selection accompanied by reductions in the level of variation observed. Expansion biased repeat instability was also observed, although the length changes recorded were all relatively modest (<15 repeats), even after 600 doublings *in vitro*. The higher levels of repeat heterogeneity detected in our system may result from either a higher intrinsic instability of the *Dmt-D* transgene, the dynamics of the repeat in the source tissues used or an effect of immortalisation. SV40 expression clearly alters progression through the cell cycle with evidence for direct effects on aspects of the DNA repair and metabolism machinery (Lanson *et al.*, 2000). Alternatively, it may simply result from the higher sensitivity of SP-PCR in detecting rarer alleles compared to the electrophoretic profiles generated using high template levels, fluorescent primers, automated fragment analysis and GeneScan software (Fortune *et al.*, 2000; Kennedy and Shelbourne, 2000).

The differences in the levels of somatic mosaicism previously reported between *Dmt-E* and *Dmt-D* transgenic mouse lines (Fortune *et al.*, 2000) remained in culture, corroborating a role for positional factors as major modifiers of trinucleotide dynamics. These results further highlight the considerable effect that flanking DNA sequences have in modifying expanded repeat stability (Brock *et al.*, 1999) and reject a global genomic destabilisation under *in vitro* growth conditions, as the cause of the trinucleotide repeat instability observed in cultured somatic cells.

Taken together, the data presented here strongly suggest that DNA replication is not the sole explanation for repeat instability, implying a role for additional *cis*- and/or *trans*-acting tissue-specific factors in the control of the dynamics of repetitive sequences in the soma. Additional evidence for non-cell division dependent instability is afforded by the age-dependent decrease in stability of maternally transmitted expanded CAG repeats in SCA1 and DRPLA transgenic mice (Kaytor *et al.*, 1997; Sato *et al.*, 1999) and the patterns of somatic mosaicism observed in various transgenic mouse models (Fortune *et al.*, 2000; Kennedy and Shelbourne, 2000; Mangiarini *et al.*, 1997; Seznec *et al.*, 2000).

In summary, cell lines derived from a transgenic mouse model for triplet repeat instability have been established and shown to retain the tissue-specific, expansion-biased instability observed *in vivo*. Under conditions of rapid cell turnover, the data described support an expansion mechanism that might not be strictly dependent on cell division, contrasting to the prevalent DNA polymerase slippage hypothesis. This readily renewable primary cell system creates novel and exciting avenues to study the complex dynamics of triplet repeats.

## 4. Investigating the role of DNA topology as a mediator of trinucleotide repeat dynamics

### 4.1. Introduction

Although multiple mutation mechanisms for trinucleotide repeats are under debate, they all share a unifying theme: the speculation that stable non-B-DNA structures, by either or both of the DNA repetitive strands, is involved in the mechanism of genetic instability, acting as the driving force for expansion (Mariappan *et al.*, 1998; McMurray, 1999; Pearson and Sinden, 1998b; Sinden *et al.*, 2002). Several pieces of evidence have suggested an association between alternative DNA structures and human disease. First, nearly all trinucleotide repeat sequences, which expand into the disease-associated range, form stable single-stranded DNA hairpins *in vitro* (Chen *et al.*, 1995; Gacy *et al.*, 1995; Mariappan *et al.*, 1996; Mitas, 1997; Zheng *et al.*, 1996), with the possible exception of single-stranded GAA and TTC oligonucleotides (Gacy *et al.*, 1998; LeProust *et al.*, 2000). Second, thermodynamic analysis of single-stranded repeats revealed that the formation of intrastrand structures, or hairpins, within Watson-Crick duplexes, is favoured for the disease-associated CAG•CTG and CGG•CCG repeats relative to the non-disease associated GTC•GAC repeats (Zheng *et al.*, 1996). Third, unusual helical properties have been identified for DNA sequences containing CAG•CTG and CGG•CCG repeats: unusual electrophoretic mobility (Chastain *et al.*, 1995; Pearson and Sinden, 1996; Pearson *et al.*, 1998b), hyperflexibility (Chastain and Sinden, 1998), and preferential nucleosome assembly have been reported (Godde and Wolffe, 1996; Wang *et al.*, 1994; Wang and Griffith, 1995). Fourth, computational analyses have confirmed that disease-associated triplet repeats generally fall into extreme categories concerning their biophysical properties, such as bendability, position preference in the DNA helix and protein-induced deformability (Baldi *et al.*, 1999). For all these reasons, it appears logical to speculate that triplet repeat instability specific to CAG•CTG, CGG•CCG, and GAA•TTC is likely to depend on the formation of unusual DNA structures.

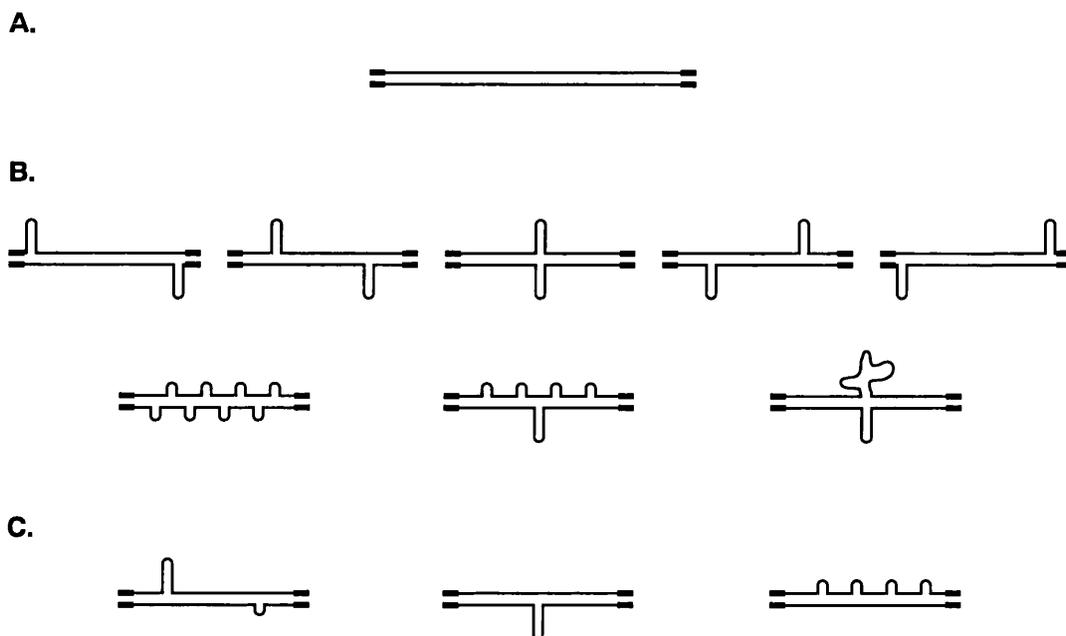
Studies carried out with single-stranded DNA sequences have revealed significant differences in the stability of the hairpins formed by individual repeats *in vitro*, which may result in different levels of instability. Among sequences able to form hairpins, the CAG•CTG trinucleotide repeats, studied in this work, have been considered to form the most stable hairpins, and folding has been reported for single-stranded DNA sequences containing as few as five or six CTG repeats, over a wide range of solution conditions (Mariappan *et al.*, 1996). It has been shown that the CTG strand forms hairpins, where the T•T mismatches are well stacked in the helix, and appear to be stabilised by two hydrogen bonds. In contrast, the A•A mismatches in the CAG strand are not well stacked in the helix, and exhibit conformational instability (Mariappan *et al.*, 1996; Mitas, 1997; Pearson and Sinden, 1998b). Additional structures have also been identified within particular tandem trinucleotides. Although single-stranded GAA repeats may not show the propensity to form

hairpins, they can form a triple helix containing non-Watson-Crick pairs (Gacy *et al.*, 1998; LeProust *et al.*, 2000; Ohshima *et al.*, 1996a), in which the third DNA strand folds into the major groove of the purine-pyrimidine duplex tract. In addition, G-rich CGG sequences also display more complex structural features. These repeats have the potential to form, not only hairpins, but also inter- or intramolecular tetraplexes (Darlow and Leach, 1998b; Mitas, 1997; Pearson and Sinden, 1998b), in which the G residues are involved in the establishment of G<sub>4</sub>-tetrads as the major stabilising force (Darlow and Leach, 1998b; Fojtik and Vorlickova, 2001; Usdin, 1998).

Double-stranded DNA sequences comprising trinucleotide repetitive tracts, have also exhibited unusual properties. Both gene fragments containing either CAG•CTG or CGG•CCG repeats, and pure duplex CAG•CTG or CGG•CCG sequences, migrate 20% faster in non-denaturing polyacrylamide gels than expected for linear duplex B-DNA (Chastain *et al.*, 1995; Chastain and Sinden, 1998). Both electrophoretic mobility analysis and cyclisation kinetic studies revealed that CAG•CTG tracts are more flexible or curved than random B-DNA (Bacolla *et al.*, 1997; Chastain and Sinden, 1998). The rapid mobility has, therefore, been suggested to result from the presence of a long flexible helical region (Chastain and Sinden, 1998). The preferential assembly of CAG•CTG repeats into nucleosomes, mentioned above, (Godde and Wolffe, 1996; Wang *et al.*, 1994; Wang and Griffith, 1995), is consistent with a curved flexible helix (Chastain and Sinden, 1998).

More interestingly, a novel form of non-B-DNA structures was identified within trinucleotide repetitive tracts. After melting and reannealing of plasmid DNA or gene fragments containing CAG•CTG or CGG•CCG repeats, a high proportion of the DNA population adopts alternative conformations that can be deduced from the retarded mobility of these new species in native polyacrylamide gels. Biochemical evidence and analysis by means of electron microscopy are consistent with the existence of slipped-stranded DNA structures formed within the triplet repeats in otherwise linear duplex molecules (Pearson *et al.*, 1997; Pearson and Sinden, 1996; Pearson *et al.*, 1998b). Two types of slipped-stranded DNA structures have been identified: homoduplex slipped structures (S-DNA) (Pearson and Sinden, 1996), which are formed between two complementary strands with the same number of repeats, paired in an out-of-register fashion; and heteroduplex slipped-stranded intermediates (SI-DNA) (Pearson *et al.*, 1997), which have different numbers of repeats in each strand (Figure 4.1). It was proposed that SI-DNA might arise through replication slippage, resulting in repeat expansions or deletions, whereas strand slippage in the absence of replication results in the generation of S-DNA. These structures are remarkably thermostable and display minimal interconversion between isomers and/or the linear duplex form under physiological conditions, possibly as a result of the formation of intrastrand hairpins, which imply that branch migration requires breaking of stabilising base pairs in the slipped-out junction and within the slipped-out hairpin (Pearson and Sinden, 1998a; Pearson and Sinden, 1998b). In S-DNA, the CAG strand is preferentially susceptible to mung bean nuclease, compared to the CTG strand (Pearson and Sinden, 1996), indicating the greater single-stranded character of the CAG strand, consistent with the lower thermal stability of the CAG hairpin relative to the CTG hairpin (Gacy *et al.*, 1995). Both the propensity for S-DNA formation and the complexity of the structures

increase with the length of the pure repeat tract (Pearson *et al.*, 1997; Pearson and Sinden, 1996) and parallels the probability of repeat expansion (Richards and Sutherland, 1992; Sinden *et al.*, 2002). Moreover, sequence interruptions within pure repeat tracts dramatically decrease the propensity to form slipped-stranded DNA structures as well as the heterogeneity of the structures formed. Thus, the total number of possible stable structures formed will be much lower for a triplet-repeat tract containing sequence interruptions, suggesting that hairpin structures containing a sequence interruption are relatively unstable compared with those formed within pure repeat tracts (Pearson *et al.*, 1998a). The protective effect of sequence interruptions in the formation of alternative DNA structures parallels an identical protective effect on the acquisition of the disease state. The interruption of the perfect CAG•CTG triplet may be significant in maintaining the stability of *SCA1* alleles (Chung *et al.*, 1993). Similarly, interruptions in expanded *FRAXA* (Eichler *et al.*, 1994) and *FRDA* alleles (Cossee *et al.*, 1997; Montermini *et al.*, 1997) clearly stabilise the repeats. Taken together these observations strongly suggest a role of slipped-stranded structures in the process of “dynamic mutation”.



**Figure 4.1. Models of S-DNA and SI-DNA structures.**

(A) Schematic representation of homoduplex B-DNA. (B) Models of various structural isomers of S-DNA generated by denaturation and renaturation of a triplet repeat tract in an out-of-register fashion. Looped-out regions can be of variable size and/or number, and they can be positioned throughout the repeat tract. (C) Models of possible SI-DNA structural isomers consisting of two single-stranded DNA sequences of different sizes. Thin lines represent trinucleotide repetitive tracts. Thick solid lines represent flanking non-repetitive sequences (Pearson and Sinden, 1996; Pearson and Sinden, 1998b).

The great biological relevance of unorthodox DNA structures is enhanced by their central involvement in hypothetical models of genetic instability, and their biophysical properties may even explain the relative stability exhibited by different repeats. A hairpin containing flexible base pairs is anticipated to melt rapidly and/or rapidly incorporate additional triplet repeats into the hairpin, relative to one containing only Watson-Crick pairs. The kinetics of melting and refolding of CTG and CAG hairpins would require fewer steps than a CGG tetraplex. Thus, on the basis of the known biophysical properties of the CCG-, CGG-, CTG- and CAG-containing single-stranded structures, it was anticipated that CTG•CAG sequences would be more prone to expansion compared with CCG•CGG sequences (Mitas, 1997). The greater expansion rate of CAG•CTG repeats in *E. coli* compared to another nine repetitive sequences supports this model (Ohshima *et al.*, 1996b). This finding corroborated the importance of DNA structural properties as central mediators of repeat metabolism.

Despite the progress achieved following a series of *in vitro* approaches, unambiguous proof of the existence of unusual DNA secondary structures *in vivo* has been extraordinarily difficult. The conformational dynamics of trinucleotide repeats was studied in *E. coli* using a bacteriophage  $\lambda$  derivative containing a long palindrome. Assuming that hairpin loop stability correlates inversely with plaque size, the measurement of plaque size would, therefore, be a reliable estimate for the stability of DNA hairpin loops in bacteria. It was suggested that both CAG•CTG and CGG•CCG sequences, containing as few as two repeat units, can fold into hairpin-type secondary structures (Darlow and Leach, 1995).

In summary, circumstantial evidence supports the idea that if DNA secondary structures exist *in vivo* they may interfere with DNA metabolism. Indeed, secondary structures may elude the cellular machinery designed to detect and repair single-stranded loops in yeast (Moore *et al.*, 1999). Nonetheless, definite proof of the existence of unconventional secondary structure in a mammalian cell environment has been particularly challenging. The possibility exists that trinucleotide repeats fold into various alternative structures that might co-exist in dynamic equilibrium *in vivo*, depending on the environmental conditions and on their interaction with cellular proteins. The possibility that DNA interactions with proteins involved in regulating DNA topology could disturb such equilibrium *in vitro* was investigated. If conformational fluctuations between structural isomers were shifted towards linear B-DNA, a stabilisation of the repeat could be expected. Using *Dmt-D* cell lines a role for DNA topology and alternative structures in the mechanism of trinucleotide repeat expansion in mammalian cells was also examined.

## 4.2. Results

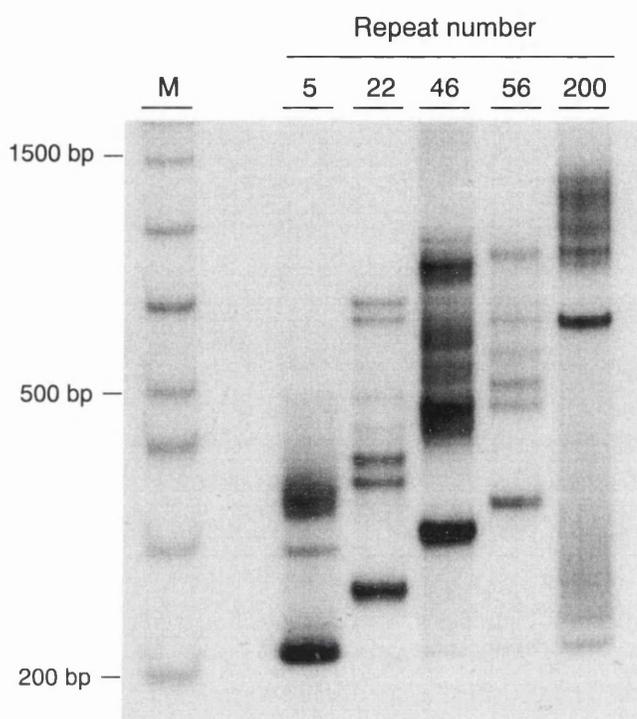
### 4.2.1. Detection of DNA alternative structures in CAG•CTG sequences generated by PCR amplification

#### 4.2.1.1. Analysis of alternative CAG•CTG-containing PCR products by native polyacrylamide gel electrophoresis

DNA sequences containing a different number of CAG•CTG repeats (ranging from five up to 200) were generated by two rounds of PCR amplification. Human genomic DNA was initially amplified by single molecule SP-PCR, with oligonucleotide primers DM-A and DM-BR (Monckton *et al.*, 1995). The PCR products obtained served as templates for a second amplification, with primers DM-A and DM-DR.

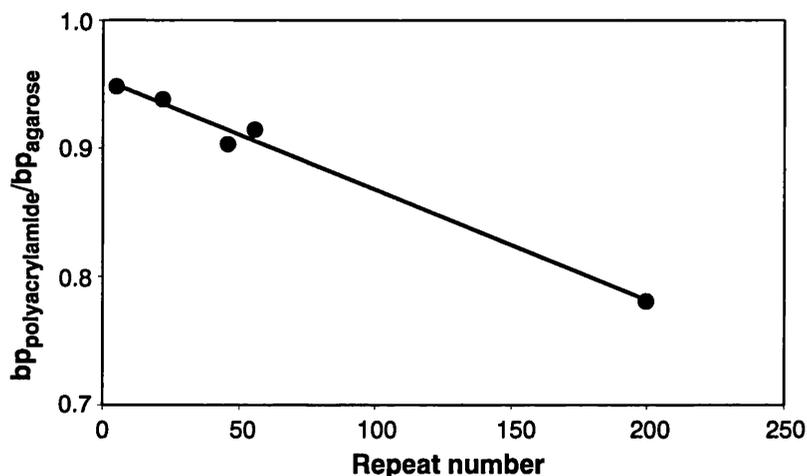
To search for the presence of alternative DNA structures in the PCR products, the amplified molecules were resolved through a native 8% (w/v) polyacrylamide gel, with 10% (v/v) glycerol (Section 2.5.7.2). Although each PCR product generated a single band following agarose gel electrophoresis in the presence of ethidium bromide (Chapter 5, Figure 5.10), complex electrophoretic profiles, consisting of a series of novel, closely spaced migrating DNA species, were observed on native polyacrylamide gels (Figure 4.2). The patterns of anomalously migrating products were very similar to those previously described for triplet repeat-containing gene fragments or plasmids following denaturation and renaturation protocols (Pearson and Sinden, 1996; Pearson *et al.*, 1998b). This observation strongly suggests that the additional bands observed correspond to alternative DNA structures, particularly S-DNA. Alternatively, given the natural tendency for *Taq* DNA polymerase slippage during DNA synthesis of trinucleotide repeat sequences *in vitro* (Lyons-Darden and Topal, 1999), it may be considered that the additional bands may also correspond to heteroduplex SI-DNA structures, consisting of different sized single-stranded DNA sequences (Pearson *et al.*, 1997). Small size DNA species were detected following amplification of a DNA sequence containing 200 CAG•CTG repeat units. The fast migrating smeary bands may be the result of non-specific DNA degradation, or they may represent SI-DNA, generated following the accumulation of large, but rare, repeat deletions during PCR amplification.

The major band, detected for each PCR product, migrated faster than expected through native polyacrylamide gels (Chastain *et al.*, 1995; Pearson and Sinden, 1996; Pearson *et al.*, 1998b). To quantify this observation, the ratio between the mobility rates through polyacrylamide and agarose gels was calculated ( $R = \text{bp}_{\text{polyacrylamide}} / \text{bp}_{\text{agarose}}$ ) (Chastain *et al.*, 1995). This calculation serves as a good estimate of the ratio between the apparent and the actual sizes for a particular band. For the major and fast migrating band this ratio varied from 0.95 for the PCR product containing 5 CTG repeats (corresponding to a ~5% increase in mobility) down to 0.78 for the longest repeat (corresponding to a ~22% increase in mobility). A linear correlation ( $R^2 = 0.9864$ ) was revealed when the  $\text{bp}_{\text{polyacrylamide}} / \text{bp}_{\text{agarose}}$  ratios were plotted against the repeat numbers (Figure 4.3), suggesting that the fast mobility of these sequences in native polyacrylamide gels is intimately

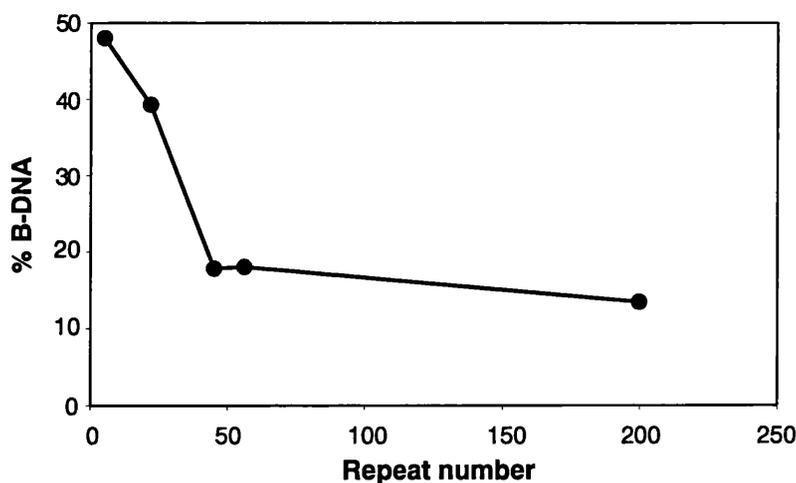


**Figure 4.2. Analysis of CAG-CTG-containing PCR products by native polyacrylamide gel electrophoresis.**

Human DNA samples derived from the *DM1* locus, carrying different repeat numbers, were amplified using oligonucleotide primers DM-A and DM-DR, and electrophoresed through a non-denaturing 8% (w/v) polyacrylamide gel, containing 10% (v/v) glycerol. The PCR products were detected by Southern blot hybridisation. The gel shows the presence of a complex pattern of multiple bands for each PCR product. The molecular size markers (M) are shown on the left.



**Figure 4.3. Increase mobility in native polyacrylamide gels as a function of repeat number.** The ratio between the relative mobility in native polyacrylamide gels and the mobility in agarose gels ( $bp_{\text{polyacrylamide}}/bp_{\text{agarose}}$ ) is plotted against the repeat number of the PCR products included in this study. The graph shows a linear relationship ( $R^2=0.9864$ ), indicating that the increase in the electrophoretic mobility through non-denaturing polyacrylamide gels increases as the CAG•CTG repeat tract lengthens.



**Figure 4.4. Percentage of B-DNA conformation as a function of repeat length.** The percentage of B-DNA detected in a non-denaturing polyacrylamide gel was estimated for each PCR product by densitometric analysis, and plotted as a function of the CAG•CTG repeat number. The graph reveals that the percentage of duplex B-DNA decreases as the trinucleotide repeat gets longer, suggesting a parallel increase in the propensity to form alternative S-DNA structures.

dependent on the repeat tract length, and may be associated with the enhanced flexibility of trinucleotide repeat tracts (Chastain and Sinden, 1998).

The additional lower mobility bands defined electrophoretic patterns that were similar for all triplet repeat lengths analysed. The slowly migrating products fell into two distinct populations with apparent size ratios ( $bp_{polyacrylamide}/bp_{agarose}$ ) of  $\sim 1.5$  and  $\sim 2.5$  (Table 4.1). It is noteworthy that the shortest PCR product carrying only five CAG•CTG repeats did not give rise to the slowest mobility band ( $R=\sim 2.5$ ), indicating that the corresponding alternative structure cannot be formed with such a low repeat number (Figure 4.2 and Table 4.1). Similar ratios of 1.25 and 2.0 have been previously reported in a study performed with plasmid DNA carrying cloned CAG•CTG repeats derived from the *DMI* locus (Pearson *et al.*, 1998b). The differences between the ratios reported here, and the ratios described by other authors, may be attributed to the mobility dependence on the length and nature of DNA sequences flanking the triplet repeat (see below).

**Table 4.1. Relative mobilities of CAG•CTG-containing PCR products through native polyacrylamide gels.**

The table displays the relative mobility of each band detected for each PCR product by native PAGE ( $bp_{poly}^a$ ), and the ratio between the relative mobility through polyacrylamide and agarose gels ( $bp_{poly}/bp_{agarose}$ ).

Band	5 CAG•CTG		22 CAG•CTG		46 CAG•CTG		56 CAG•CTG		200 CAG•CTG	
	$bp_{poly}$	$bp_{poly}/bp_{agarose}$								
1	214	0.95	260	0.94	315	0.90	346	0.91	633	0.78
2	297	1.31	373	1.35	465	1.33	478	1.26	1048	1.29
3	352	1.56	403	1.45	534	1.53	517	1.37		
4			447	1.62	602	1.72	572	1.51		
5			494	1.78	837	2.40	639	1.68		
6			640	2.31			894	2.36		
7			708	2.56						

<sup>a</sup> poly, polyacrylamide

Assuming that the fast mobility band represents linear duplex B-DNA (Pearson and Sinden, 1996; Pearson *et al.*, 1998b), a densitometric analysis was performed in an attempt to quantify the propensity of different repeat sizes to adopt alternative S-DNA. The relative intensity of the highest mobility band was determined and plotted as a function of the repeat number (Figure 4.4). The graph shows that the proportion of linear duplex DNA decreases as the repeat gets longer, reaching a plateau around 50 CAG•CTG repeat units. This observation indicates that lengthening of the repeat tract enhances the propensity to form S-DNA structures, as described in earlier reports (Pearson *et al.*, 1998b).

It might be speculated that, for the same repeat number, the tendency for out-of-register reannealing and S-DNA formation, increases with the size of non-repetitive flanking DNA sequences. Longer non-repetitive DNA sequences, flanking the CAG•CTG repeat, would reanneal more readily than shorter flanking sequences, creating greater opportunities for the trinucleotide repetitive tract to reanneal out-of-register, forming intrastrand loops and/or hairpins. For the same

repeat number, longer non-repetitive flanking sequences would therefore be associated with lower percentages of conventional homoduplex B-DNA, and a greater extent of S-DNA formation. Indeed, for long CAG•CTG repeat tracts (>50 units), whereas a flanking region of 211 bp is associated with ~15% of B-DNA (Figure 4.4), a flanking sequence of 113 bp is associated with ~30% of B-DNA (Pearson *et al.*, 1998b). Similarly, different sized non-repetitive flanking DNA sequences, may explain the greater  $\text{bp}_{\text{polyacrylamide}}/\text{bp}_{\text{agarose}}$  ratios determined in this study, relative to the lower ratios reported elsewhere (Pearson *et al.*, 1998b), again suggesting that longer flanking DNA sequences may facilitate the formation of alternative structures upon reannealing.

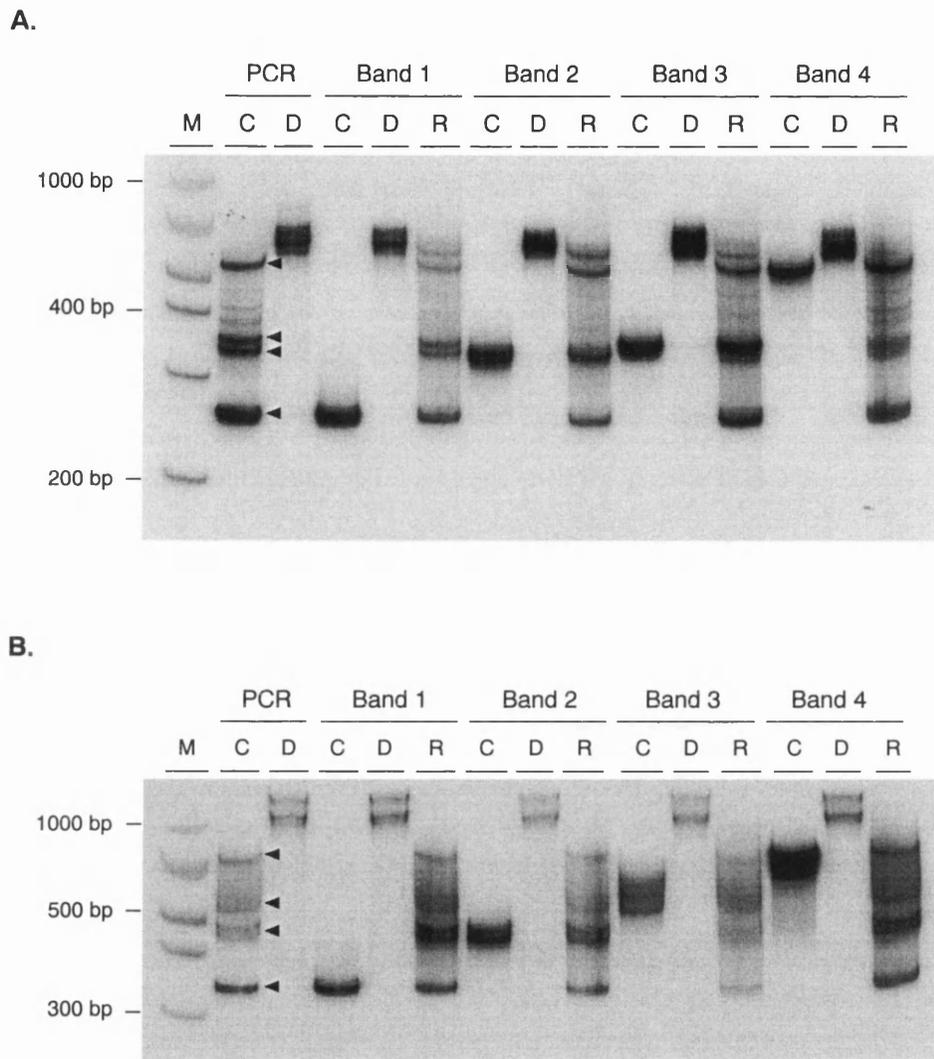
#### 4.2.1.2. Confirmation of S-DNA folding within CAG•CTG sequences

The similarity between the electrophoretic profiles observed when PCR products were electrophoresed in native polyacrylamide gels and previous reports (Pearson and Sinden, 1996; Pearson *et al.*, 1998b), was considered as a strong indication that slipped-stranded non-B-DNA structures were already present in the PCR products generated under standard conditions. Further experimental support was obtained by melting and reannealing DNA purified from single bands in non-denaturing polyacrylamide gels.

Individual bands (Figure 4.5, black arrowheads) were isolated from two PCR products containing either 22 or 56 CAG•CTG repeats. These particular PCR products were selected to perform this experiment, since they gave rise to discrete sharp individual bands, easily purified from the gel (Figure 4.5, black arrowheads). DNA was eluted from the gel fragments by simple diffusion and then ethanol precipitated, with the addition of linear polyacrylamide as a DNA carrier (Section 2.5.7.3). For each DNA sample derived from a single band, two treatments were performed. The denaturation procedure consisted of heating the DNA at 100°C for five minutes in the presence of 40% (v/v) formamide, followed by rapid chilling on ice. The reannealing protocol consisted of a first step of DNA melting, performed at 100°C for five minutes, followed by slow cooling to room temperature, over a period of at least three hours.

The samples were analysed by PAGE (Figure 4.5). All gel-purified DNA samples migrated as single bands (Figure 4.5, lanes C), confirming their high stability throughout the elution and precipitation procedures. DNA melting in the presence of formamide resulted in the detection of a couple of very low mobility bands (Figure 4.5, lanes D). These bands are believed to correspond to slowly migrating linear single-stranded DNA. The full set of anomalously migrating bands, identical to that observed for the original PCR product, was restored following the reannealing protocol (Figure 4.5, lanes R). The results were identical for both PCR products included in this study, independently of the number of repeats carried by the repetitive sequence.

The observation that the reannealing protocol resulted in a pattern of products that was very similar to that previously observed in the original PCR products, further confirmed that the slow migrating DNA species detected correspond to slipped-stranded DNA (S-DNA).



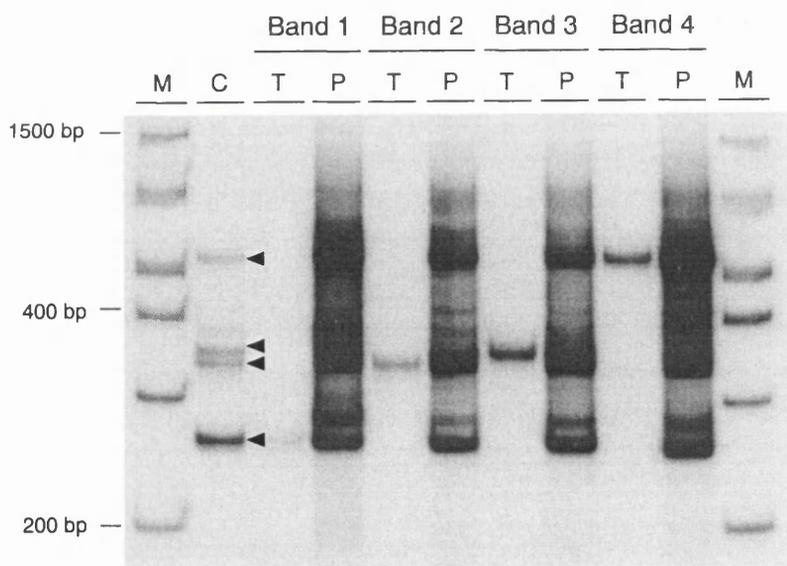
**Figure 4.5. Reannealing experiments on PCR products containing CAG-CTG repeats.**

PCR products containing either (A) 22 or (B) 56 CAG-CTG repeats were amplified with oligonucleotide primers DM-A and DM-DR. DNA was extracted from single bands following native PAGE (Bands 1-4, indicated by black arrowheads, ▲) and re-electrophoresed (lanes C). Duplex DNA samples were melted at 100°C in the presence of 40% (v/v) formamide, and originated a couple of presumed single-stranded DNA bands of low mobility (lanes D). Alternatively, DNA samples were denatured at 100°C, and subsequently reannealed by slow decrease of temperature for over 3 hours (lanes R), reoriginating a complex pattern of bands, previously observed in the original PCR product. The molecular size markers (M) are shown on the left.

#### 4.2.1.3. Generation of alternative structures by PCR amplification

If the additional bands observed do represent S-DNA conformations adopted by CAG•CTG repetitive structures, these results indicate that non-B-DNA structures are generated under standard PCR cycling conditions, since they were previously detected in the original amplified products (Figure 4.2). To confirm that PCR amplification is sufficient to produce alternative DNA structures, gel-purified DNA samples derived from single bands (Figure 4.6, black arrowheads) and carrying 22 CAG•CTG repeats, were used as template in a second amplification reaction with the same oligonucleotide primers. The DNA samples were analysed by native PAGE before and after the second round of PCR amplification (Figure 4.6). The products generated by the second amplification reaction migrated as a complex pattern of heterogeneous bands, exhibiting electrophoretic profiles very similar to those detected in the original PCR product, as well as following the reannealing protocol, described in Section 4.2.1.2.

In summary, the amplification of DNA samples, derived from the human *DMI* locus by standard PCR techniques generates a complex mixture of alternative DNA conformations, which are interconvertible following a reannealing protocol, and most likely to be slipped-stranded DNA structures.



**Figure 4.6. Generation of alternative DNA structures by PCR amplification.**

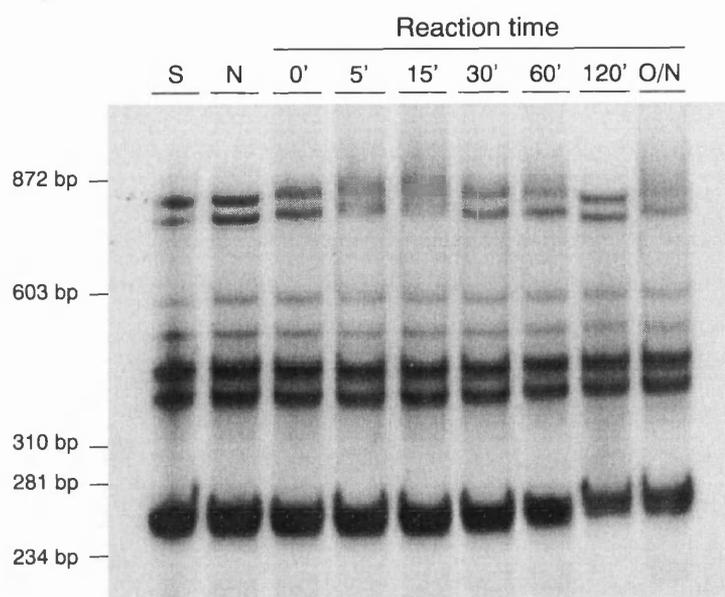
DNA samples, containing 22 CAG•CTG repeat units, were eluted from four bands (Bands 1-4, indicated by black arrowheads, ◄) cut out from a native 8% (w/v) polyacrylamide gel and used as templates for a second PCR amplification. Each DNA sample was re-electrophoresed before (T) and after the second PCR amplification (P). The amplification resulted in a complex pattern of bands very similar to that previously observed in the original PCR product (C) from which the bands were eluted. The molecular size markers (M) are shown on the left.

## 4.2.2. Testing RecA activity on CAG•CTG repetitive sequences

*In vitro* studies indicate that in the presence of ATP, RecA protein promotes the strand exchange of single-strand DNA fragments with homologous duplex DNA. The reaction has three distinct steps: (a) RecA protein polymerises on the single-stranded DNA, (b) the nucleoprotein filament binds the duplex DNA and searches for a homologous region, (c) the strands are exchanged (Radding, 1991). More interestingly, RecA has been reported to recombine short stable hairpin substrates sharing homologous stems through a four-strand exchange model that requires partial unwinding of DNA hairpins (Gamper *et al.*, 2000). Furthermore, RecA is also able to catalyse the realignment of stochastically paired suboptimal frames in order to maximise the aligned register. RecA binds a single-stranded oligonucleotide and mediates such realignments of suboptimally paired frames in a slow, time- and ATP-dependent manner, probably by sliding the paired strand across repetitive DNA sequences (Sen *et al.*, 2000). RecA was consequently considered as a good candidate to bind and mediate triplet repeat metabolism, therefore, its ability to interact with CAG•CTG sequences, to promote strand exchange and/or to interconvert alternative slipped structures was assessed *in vitro*.

### 4.2.2.1. RecA activity assay on CAG•CTG repetitive tracts

Making use of the PCR products previously generated, a simple RecA activity assay (Bennett and Holloman, 2001) was carried out on a PCR product containing 22 CAG•CTG repeat units, in order to test the ability of the protein to catalyse strand exchange reactions, and interconvert alternative S-DNA structures formed by triplet repeats. The assay was based on the principle that RecA-catalysed interconversion between structural isomers would alter the electrophoretic profile of the reaction product in a native polyacrylamide gel. Briefly, duplex DNA molecules were incubated in the presence of a melted single-stranded (CTG)<sub>12</sub> oligonucleotide, RecA and ATP for increasing periods of time (up to 16 hours; for further details see legend of Figure 4.7). Reactions were stopped at different time points by the addition of EDTA, which caused depletion of free magnesium ions in solution, an essential co-factor for RecA catalysis. Analysis of the reaction products by native PAGE and subsequent Southern blot hybridisation, failed to reveal differential DNA electrophoretic profiles at any time point, even after 16 hours of incubation with RecA in the presence of ATP. All individual bands, corresponding to a heterogeneous population of structural isomers present in the original substrate, remained unaffected throughout the reaction (Figure 4.7, lane S). Furthermore, with the exception of subtle and presumably not significant changes in the relative mobility of the fast migrating band after 120-minute and overnight incubations, no DNA species showed major and consistent gel retardation or band shift, indicating that RecA failed to bind duplex alternative DNA structures formed within a CAG•CTG repetitive sequence containing 22 repeat units. In summary, RecA appears to be unable to interconvert alternative S-DNA structures adopted *in vitro* by CAG•CTG PCR products, containing as few as 22 repeats, at least under the conditions tested in this assay.



**Figure 4.7. Assay for RecA activity on CAG•CTG repetitive sequences.**

Reactions (20  $\mu$ l) containing 16  $\mu$ M single-stranded (CTG)<sub>12</sub> oligonucleotide (previously melted at 100°C for 5 minutes and rapidly cooled down on ice) in 30 mM tris-acetate, pH 8.0, 1 mM magnesium acetate, 2 mM DTT, 100  $\mu$ g/ml BSA, 1.3 mM ATP, 20 mM phosphocreatine, 10 units/ml phosphocreatine kinase and 1  $\mu$ M RecA protein were preincubated at 37°C. After 10 minutes, magnesium acetate was increased to 13 mM and duplex PCR product containing 22 CAG•CTG repeats (amplified with DM-A and DM-DR) was added to a final concentration of 8  $\mu$ M. Thus, the molar ratio of single-stranded DNA to duplex DNA was 2:1. After further incubation up to a maximum of 16 hours, reactions were quenched by the addition of 15 mM EDTA and rapidly frozen on a mixture of dry ice and absolute ethanol. Reaction products, collected at the time points indicated in the figure, were electrophoresed on a native 8% (w/v) polyacrylamide gel, transferred onto a nylon membrane by Southern “squash” blot and detected by hybridisation with a radio-labelled DM56 probe. Two additional controls were included, the original duplex DNA substrate (S), and a reaction mixture incubated for 16 hours in the absence of RecA (N). No obvious differences in the electrophoretic profile were observed at any time point. The molecular weight markers are indicated on the left.

It is important to be aware of the lack of a positive control for RecA activity in this study. Despite the use of standard reaction conditions, the possibility remains that RecA was not fully functional and therefore unable to catalyse the strand exchange reaction. In addition, the assay was designed to assess the effect of RecA activity on a possible dynamic equilibrium between alternative DNA structures formed by CAG•CTG-containing PCR products. This hypothetical equilibrium might have been disturbed by RecA, but given its dynamic nature, the outcome could not be detected by our assay. Ideally, RecA activity should be individually tested on DNA species purified from single bands.

In conclusion, RecA strand exchange activity was not detected on short trinucleotide repetitive sequences. However, given the highly preliminary nature and the limitations of the assay, the results presented must be considered inconclusive. Rather than being performed on a mixture of putative alternative DNA conformations, the analysis should be extended to individual DNA structural isomers, and possibly to longer repeat tracts.

### **4.2.3. *Novobiocin as a genotoxic modifier of trinucleotide repeat instability: a possible role for topoisomerase II in triplet repeat metabolism***

Some of the physical properties of CAG•CTG repeats are mediated by topological features associated with their increased flexibility (Chastain and Sinden, 1998). Being more flexible and more highly writhed than random B-DNA, CAG•CTG repeats are expected to act as sinks for the accumulation of super-helical density (Bacolla *et al.*, 1997). Statistical mechanical calculations revealed that free energies of supercoiling are lower for trinucleotide repeats than for B-DNA (Gellibolian *et al.*, 1997), which led to the suggestion that trinucleotide repeat instability might be mediated by the accumulation of supercoiling within the repeats. Given the biophysical properties of expanded repeats and their propensity to adopt unusual non-B-DNA structures *in vitro*, it may be hypothesised that proteins involved in resolving complex DNA structures might be important in the mutational process. The formation of alternative DNA conformations, such as S-DNA, in duplex chromosomal DNA would be predicted to alter the normal displacement of superhelical tension. It is not unreasonable to imagine that proteins that regulate DNA superhelical tension may play a role in trinucleotide repeat metabolism.

The various problems of disentangling DNA strands or duplexes in a cell are all rooted in the double-helical structure of DNA. Torsional constraints are introduced during replication and transcription, due to underwinding and overwinding of the helices, causing a dramatic alteration in the local superhelical densities of DNA (Wang, 1996). DNA topoisomerases solve the topological problems associated with DNA replication, transcription, recombination and chromatin remodelling by either passing one strand of the DNA through a transient break in the opposing strand (type I subfamily), or by passing a region of duplex from the same or a different molecule through a double-stranded gap generated in a DNA molecule (type II subfamily). The primary cellular functions of these enzymes include their roles in replication, transcription, chromosome condensation and maintenance of genome stability (Nitiss, 1998; Wang, 1996). Eukaryotic topoisomerase type I plays a major role in supporting fork movement during replication and in relaxing transcription-related supercoils, and is indispensable during development and probably also during cell division (Champoux, 2001; Wang, 1996). Topoisomerase II is responsible for unlinking intertwined daughter complexes during DNA replication, it also contributes to DNA relaxation during transcription and it is involved in the final stage of chromosome condensation, and probably in certain phases of decondensation (Champoux, 2001; Wang, 1996). Both types of DNA topoisomerases may also play a role as suppressors of mitotic recombination (Wang, 1996). In addition, these enzymes adjust the steady-state level of DNA supercoiling, both to facilitate protein interactions with the DNA and to prevent excessive supercoiling that might be deleterious for the cell (Berger, 1998; Champoux, 2001; Wang, 1996). Interestingly, topoisomerase II is capable of recognising and interacting with secondary structures within nucleic acids, such as hairpins (Froelich-Ammon *et al.*, 1994), and it preferentially binds to DNA containing mismatches (Bigioni *et al.*, 1996). This phenomenon might actually be further amplified through the formation

of molecular complexes between topoisomerases and proteins involved in DNA mismatch repair (Larsen and Skladanowski, 1998). As a result, it appears logical to speculate that topoisomerase II represents a plausible candidate protein in the mutational pathway of trinucleotide repeats.

A wide variety of topoisomerase-targeted drugs have been identified. “Topoisomerase poisons” generate cytotoxic lesions by trapping the enzymes in covalent complexes on the DNA and thereby stabilising covalent complexes that are normally transient intermediates in the catalytic cycle of the enzyme (Burden and Osheroff, 1998; Capranico and Binaschi, 1998), thus generating double strand breaks that may lead to cell death (Kaufmann, 1998). These topoisomerase poisons include both antimicrobials (Hooper, 1998), but also anticancer chemotherapeutics, some of which are currently in widespread clinical use (Burden and Osheroff, 1998). Topoisomerase “catalytic inhibitors”, unlike topoisomerase poisons, are thought to inhibit specific steps of the catalytic cycle, without stabilising covalent intermediates (Andoh and Ishida, 1998). Novobiocin is a non-specific catalytic inhibitor of the ATPase activity of type II DNA topoisomerases (Lewis *et al.*, 1996). The accumulation of positive supercoils in a plasmid was observed when an *E. coli* strain, permeable to novobiocin, was treated with the drug (Lockshon and Morris, 1983). The effect of novobiocin on trinucleotide repeat dynamics was consequently tested on *Dmt-D* mouse cells.

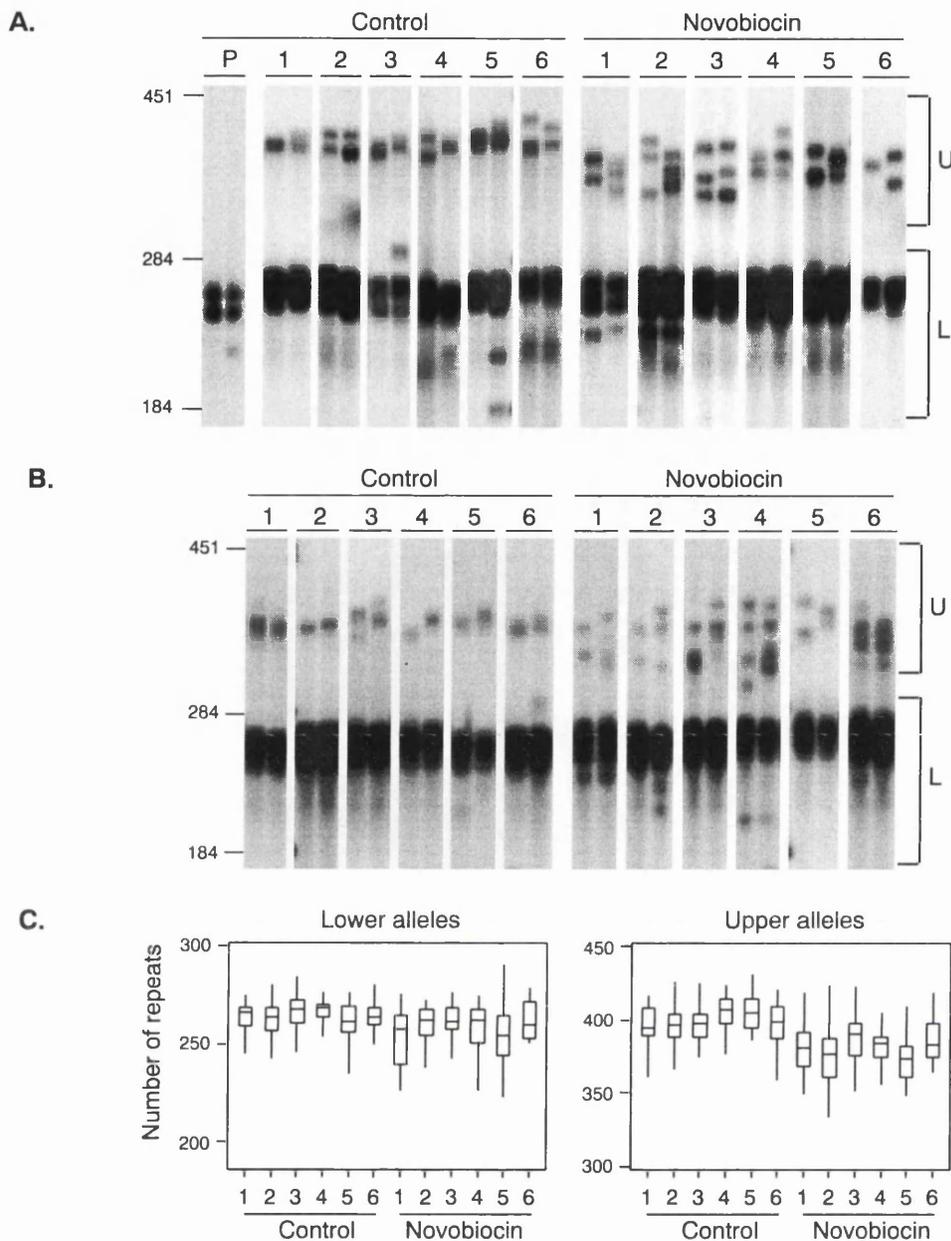
#### 4.2.3.1. D2763K cell line

Transgenic *Dmt-D* mouse kidney cells (line D2763K, Table 3.2) were treated with novobiocin, a potent inhibitor of DNA topoisomerase type II. Six replicate cultures were exposed to 60  $\mu$ M of novobiocin for 73 days (60 population doublings) maintained in parallel with six untreated controls (77 days, 63 population doublings). At the end of the exposure period the degree of repeat length variation was assessed in all cultures using sensitive SP-PCR procedures and compared to that observed in the starting culture. High input DNA analyses revealed that during the experimental period the cell line had split into two subpopulations, here described as “upper” (U,  $\geq 300$  repeats) and “lower” (L,  $< 300$  repeats). More interestingly a decreased rate of expansion was observed in novobiocin treated cells, not only at the end of the treatment, but also at an intermediate point, when novobiocin treated and control cells had undergone 36 and 39 population doublings, corresponding to 39 and 47 days in culture, respectively (Figures 4.8.A and 4.8.B).

Single molecule analysis was performed, and 50-100 transgene molecules collected at the end of the treatment were accurately sized for each individual subpopulation (L or U) of cells, in order to quantify the effect of novobiocin on the repeat dynamics in cultured mouse cells.<sup>4</sup> Median rates of expansion, corrected for time and population doublings, were calculated for each subpopulation in treated and control cultures, and compared using the two-tailed Mann-Whitney *U* test (Figure 4.8.C). Analysis of the control cultures revealed that in population L the repeat had expanded only slowly (median gain = 0.043 repeats per day); in population U however, the repeat expanded rapidly (median gain = 1.78 repeats per day). Both populations of cells were also

---

<sup>4</sup> Some of the single molecule SP-PCR amplifications were performed by Sanam Mustafa, a summer student working under my supervision.



**Figure 4.8. Novobiocin treatment and expanded CAG-CTG repeat dynamics in D2763K cells.**

The autoradiographs shown are representative SP-PCR analyses of DNA samples extracted from six replicate D2763K cultures treated with 60  $\mu$ M novobiocin. Control cultures maintained for approximately the same amount of time and the progenitor culture (P) from which all cells were derived at day zero are also shown. Fresh media with or without novobiocin was added to the cultures every 2-3 days and the cells were passaged weekly. Repeat size variability was monitored (**A**) at the end of the treatment (73 days of exposure to novobiocin, 77 days in culture for the controls) and (**B**) at an intermediate time point (47 days of exposure to novobiocin, 39 days in culture for the controls). Note that the D2763K cultures spontaneously split into two main populations of cells labelled "upper" (U) and "lower" (L). An average of five to 50 DNA molecules were amplified in the reactions shown. The scale on the left displays the molecular weight markers converted into CTG repeat numbers. (**C**) The boxplots show the degree of variation observed in both the U and L populations in treated and control cells at the end of the treatment. The top and bottom of the boxes correspond to the third (Q3) and first quartiles (Q1), respectively and the line across the box displays the median repeat number. The lines extending from the top and the bottom of the boxes, include values that fall inside the lower and upper limits:  $Q1-1.5(Q3-Q1)$  and  $Q3+1.5(Q3-Q1)$ , respectively.

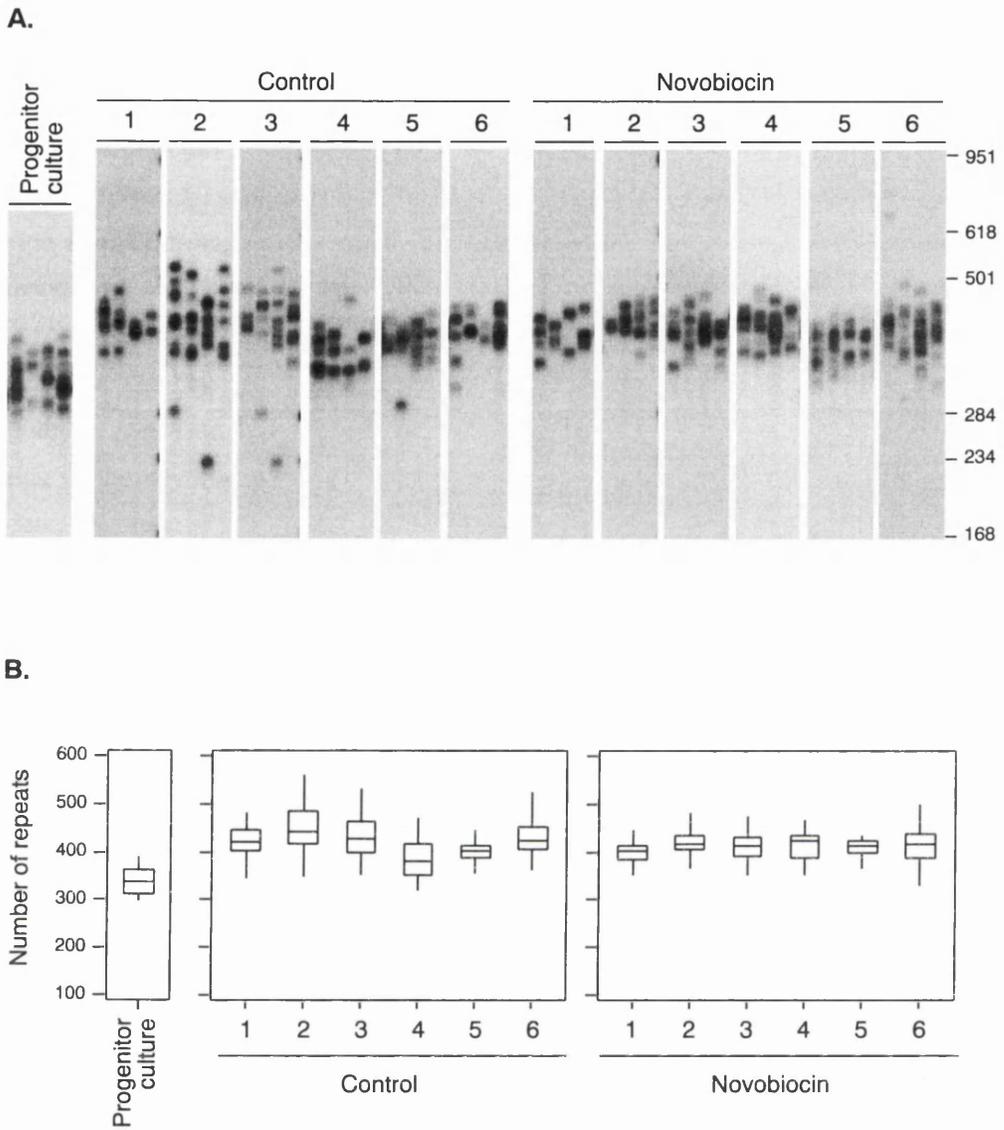
observed in the treated cultures in similar proportions. Although most dramatic in population U, the rate of expansion in both populations of treated cells was significantly reduced (median gain = 1.65 repeats per day for population U,  $p=0.0082$ ; and  $-0.017$  repeats per day for population L,  $p=0.0131$ , two tailed Mann-Whitney  $U$  test). The differences between treated and control cultures were also found to be significant when the median repeat gain per population doubling was compared between cultures. In controls, the median repeat expansion was 0.052 repeats per population doubling for population L and 2.17 for population U, whereas novobiocin treated cells showed median repeat gains of  $-0.017$  per population doubling for population L, and 2.00 for population U ( $p=0.0131$  for populations L and  $p=0.0082$  for population U, two-tailed Mann-Whitney  $U$  test). These data strongly suggested that inhibition of topoisomerase II resulted in a decrease in the rate of expansion, possibly mediated by alternative processing of mutation intermediates.

#### 4.2.3.2. D2763Kc2 cell line

Encouraged by these results we sought to determine if this effect could be replicated in other cell lines. However, rather than using the D2763K cell line with the two subpopulations, we used a cloned derivative in which the repeat expands rapidly with time (D2763Kc2, Figure 3.8). Furthermore, since the D2763Kc2 cell line was derived from a single cell by limiting dilution, it is expected to exhibit reduced genetic variability between individual cells when compared to the progenitor culture, which might have accumulated multiple genetic mutations prior to immortalisation, giving rise to a heterogeneous genetic background within the cell population.

D2763Kc2 cells were exposed to 60  $\mu\text{M}$  of novobiocin for 99 days and 80 population doublings. Control replicates were grown for the same number of population doublings over a period of 97 days. Novobiocin revealed a minor effect on the growth rate of D2763Kc2 treated cells, causing a 6% decrease in the population doubling time. Previous studies indicated that cell proliferation rates were not the most critical factor underlying differences in expansion rates between cell lines (Chapter 3). Nevertheless, the median repeat change was corrected for both time in culture and number of population doublings and compared between controls and treated cells.

SP-PCR amplification of an average number of five to 30 molecules per reaction revealed that the transgenic repeat length continued to expand rapidly in the control cells, but failed to detect any major differences in the levels of triplet repeat instability between control and treated cells (Figure 4.9.A). Single molecule analyses allowed the quantification of the expansion rates in cultured cells (Figure 4.9.B) and confirmed that the transgenic repeat length expanded rapidly in the controls (median rate of expansion = 0.792 repeats per day after 97 days). As in the previous experiment, novobiocin treatment did result in a decrease in the median rate of expansion (median rate of expansion = 0.696 repeats per day), however the measured reduction was not statistically significant for the number of molecules individually sized ( $p=0.5752$ , two-tailed Mann-Whitney  $U$  test).



**Figure 4.9. Novobiocin treatment and expanded CAG-CTG repeat dynamics in D2763Kc2 cells.**

(A) The autoradiographs shown are representative SP-PCR analyses of DNA samples extracted from replicate D2763Kc2 cells cultured for 80 population doublings with 60  $\mu$ M novobiocin (99 days), control cultures maintained for 82 population doublings (97 days), and the progenitor culture from which all cells were derived at day zero. The D2763Kc2 cell line was cloned by limiting dilution from D2763K (see Figure 3.8) and the cells passaged and analysed as indicated in the legend to Figure 4.8. An average of five to 30 transgene molecules were amplified in independent reactions. The molecular weight markers were converted into CTG repeat number and shown on the right. (B) The boxplots show the degree of variation observed in treated and control cultures as described in Figure 4.8. Differences between median rates of expansion, corrected for time and population doublings, were not significantly different according to a two-tailed Mann-Whitney *U* test.

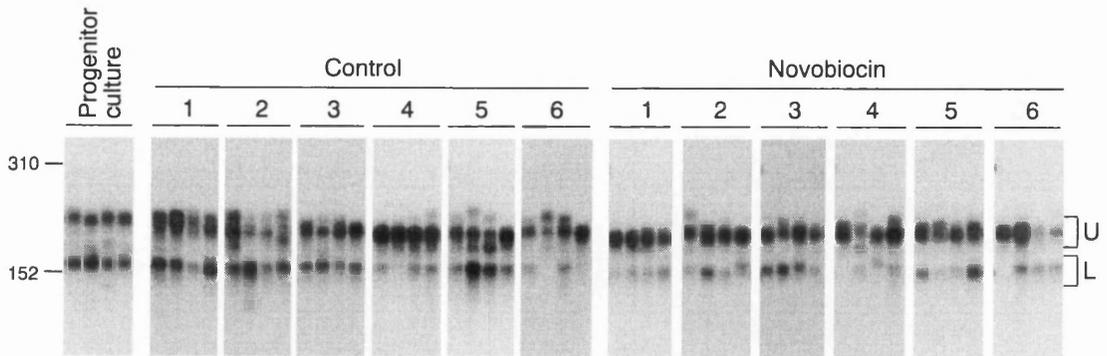
#### 4.2.3.3. D4132K cell line

Both D2763K and D2763Kc2 cell lines had already undergone a considerable number of population doublings prior to novobiocin treatment. When the treatment was initiated D2763K cells had been growing for 263 days (131 population doublings), while D2763Kc2 cells had been cultured for 186 days (45 population doublings) since the establishment of the original culture from the progenitor tissue. Given that trinucleotide repeat dynamics might be more faithfully replicated in cultured *Dmt-D* mouse cells over the first passages (Chapter 3), the effect of novobiocin on repeat instability was tested on a third cell line, which had been growing *in vitro* for a shorter period of time, closely resembling a primary mouse cell culture, prior to spontaneous immortalisation. The D4132K cell line was established from a kidney tissue sample harvested from a six-month-old *Dmt-D* mouse, and cultured for 84 days and only 12 population doublings, prior to the starting of the novobiocin treatment. By this point D4132K cells already exhibited a constant growth rate, characterised by a population doubling time of 60 hours, considerably higher than those determined for both D2763K and D2763Kc2 cell lines, consistent with the lack of spontaneous immortalisation. Six replicates were treated with 60  $\mu$ M of novobiocin for 86 days (34 population doublings) and the repeat dynamics assessed as described.

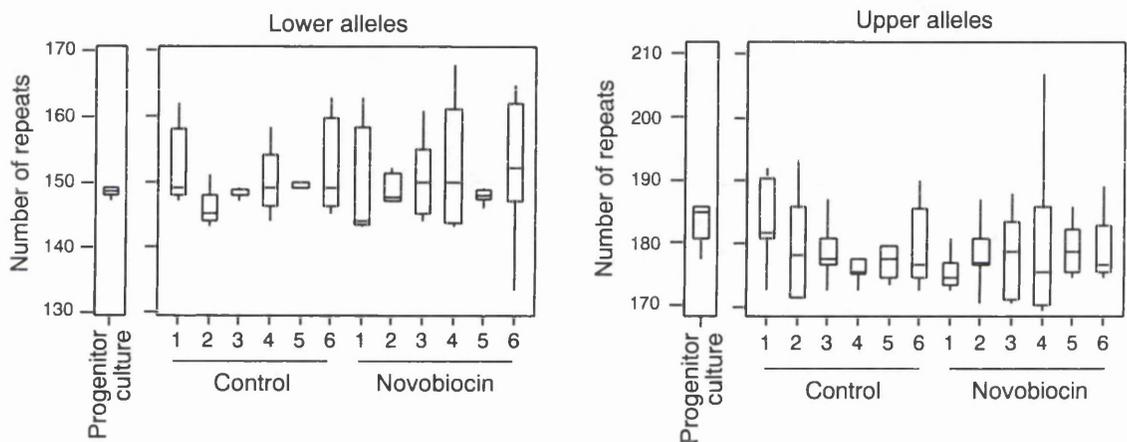
As observed with D2763K cells, SP-PCR analyses of transgene molecules extracted from D4132K cells also revealed that the cell line had split into two subpopulations: “lower” population (L, <170 repeats) and “upper” population (U,  $\geq$ 170 repeats). The establishment of the two distinct populations happened early during the establishment of the culture, since a bimodal profile of allele sizes was already detected in the progenitor culture (Figure 4.10.A). The results also show that the transgenic repeats did not expand rapidly in this cell line, neither in the control cells nor in the novobiocin treated cultures. In fact, the longer alleles appear to have contracted during the course of this experiment.

Single molecule SP-PCR analyses were carried out on a low number of “lower” (L) and “upper” (U) alleles (ten to 20 molecules), in order to quantify the median repeat change rate (Figure 4.10.B). Overall, when the two subpopulations were not considered individually, both control and treated cultures exhibited modest levels of repeat contraction during the course of this experiment, with median repeat changes of -0.0279 and -0.0343 repeats per day, respectively. As a result, this cell line may not be a powerful tool to test the stabilising effect of novobiocin on expanded trinucleotide repetitive tracts. When lower alleles were considered individually, a median expansion of 0.0626 repeats per day was determined for control cells, whereas the same alleles only gained 0.0347 repeats per day in novobiocin treated cultures. Nevertheless, the difference between the rates of expansion was not statistically significant ( $p=0.8102$  for lower alleles, two-tailed Mann-Whitney *U* test). Despite their larger size, longer alleles underwent overall contraction in both cultures, exhibiting median repeat number changes of -0.0803 and -0.0880 repeats per day in control and novobiocin treated cells, respectively. Yet again, the difference was not statistically significant ( $p=0.6889$  for upper alleles, two-tailed Mann-Whitney *U* test). Novobiocin might have

**A.**



**B.**



**Figure 4.10. Novobiocin treatment and expanded CAG-CTG repeat dynamics in D4132K cells.**

(A) The autoradiographs show representative SP-PCR amplifications of DNA samples extracted from replicate D4132K cell cultures established from a kidney tissue sample harvested from a six-month-old *Dmt-D* mouse. Treated cells were cultured in the presence of 60  $\mu$ M novobiocin for 86 days (34 population doublings). Control cells were grown for the same period of time and population doublings in the absence of the drug. The cells were passaged and analysed as indicated in the legend to Figure 4.8. Two main subpopulations of alleles were detected in culture and named “lower” (L) and “upper” (U) as previously. An average of ten to 20 molecules were amplified in each reaction. The molecular weight markers converted into CTG repeat numbers are displayed on the left. (B) The boxplots show the degree of variation observed in treated and control cultures as described in Figure 4.8. Median rates of expansion were compared between treated and control cultures and found not to be significantly different.

restricted the expansion-biased repeat instability, but the differences were too subtle to be statistically significant, at least for the low number of molecules individually sized.

In summary, whereas topoisomerase type II inhibition by novobiocin showed a statistically significant stabilising effect on the trinucleotide repeats dynamics in D2763K cells, the results obtained with two additional cell lines may suggest a similar but extremely subtle effect, yet lacking statistical significance. A higher number of transgene molecules, collected from D2763Kc2 and D4132K cells, should be individually sized by single molecule SP-PCR, in order to clarify these inconclusive results.

### 4.3. Discussion

The molecular mechanisms responsible for the genetic instability observed in triplet repeat disorders are likely to involve the unique structural properties associated with simple trinucleotide sequences, such as the formation of flexible hairpin structures, which may interfere with the activity of enzymes involved in DNA replication, transcription, repair and/or recombination. Both genetic and biochemical data suggest that trinucleotide repeat instability is a replication/repair error that is dependent on improper DNA secondary structure formation at the repeating region of the affected gene (McMurray, 1999; Richards and Sutherland, 1994; Wells, 1996). The formation of DNA structures such as hairpins or slipped-stranded structures might therefore be responsible for the unusual biology associated with trinucleotide repeats.

Alternative non-B-DNA conformations, presumably S-DNA structures, were generated by standard PCR amplification of human DNA samples. The PCR products migrated in native polyacrylamide gels as a broad distribution of distinct products ranging up to more than twice their actual size. The reduced gel mobility of the additional bands observed is consistent with that expected for DNA containing bends introduced from the three- and four-way junctions generated by the looped-out strands (Pearson and Sinden, 1998b). Mimicking single-stranded DNA reannealing *in vitro*, by denaturing and renaturing DNA sequences, presumably with equal number of repeats, confirmed that the DNA species detected in native polyacrylamide gels consisted of stable alternative DNA conformations, formed within the CAG•CTG tracts, previously described as slipped-stranded structures, S-DNA. (Pearson and Sinden, 1996). However, it must be also considered that DNA polymerase slippage might have occurred during PCR amplification, originating heteroduplex double-stranded DNA products. Therefore, at least some of the DNA species detected by native PAGE may indeed correspond to SI-DNA structures, formed by the reannealing of single-stranded trinucleotide sequences with different repeat numbers (Pearson *et al.*, 1997).

The very nature of the repeating units, which can form multiple out-of-register mispairings, involving different lengths of loop-outs at multiple locations throughout the repeat tract, leads to the generation of a high degree of structural variability. The complex pattern of bands observed for each PCR product is believed to reflect multiple structural polymorphisms within a run of triplet

repeats. The complexity of the S-DNA populations, as well as the propensity for folding into S-DNA structures was shown to increase with increasing length of the repeat tract, corroborating the slipped nature of the DNA structures formed by CAG•CTG-containing PCR products. Previous electron microscopy analyses (Pearson *et al.*, 1998b) confirmed that S-DNA secondary structures occur at random locations throughout the repeat tract, and consist of a heterogeneous population of products. Nevertheless, the various structural isomers detected did not result in an evenly distributed smear of products: certain structural variants were favourable, resulting in major alternative structures observed as major bands, as previously reported (Pearson and Sinden, 1996; Pearson *et al.*, 1998b).

Interestingly, short double-stranded repeat sequences, containing repeat lengths that are not associated with disease (5 and 22 CAG•CTG repeat units), also exhibited the potential to generate stable alternative structures *in vitro*. The ability for S-DNA formation *per se* does not distinguish long unstable repeats from short stable repeats. Therefore, S-DNA formation does not entirely explain why long repeats, but not short repeats, expand. Longer disease-associated repeats, however, displayed a higher propensity to form S-DNA structures, as estimated by the percentage of linear B-DNA detected by native PAGE. Furthermore, it was previously speculated that despite their ability to form stable non-B-DNA structures, short repeats are not prone to expansion given the short lifetime of the structures formed, in contrast with long repeat stretches, which form alternative structures with notably longer lifetimes (Gacy and McMurray, 1998). Therefore, the ability to form DNA secondary structures might be necessary, but not sufficient, to explain the molecular mechanisms underlying trinucleotide repeat expansion in human disease. Quantitative differences in the propensity for alternative folding and/or in the kinetic properties may explain why expansion occurs with higher frequency at long repeats, but not at short repeats.

Alternative stable DNA loops containing triplet repeats are inefficiently repaired by FEN1/rad27 both *in vitro* (Henricksen *et al.*, 2000; Spiro *et al.*, 1999) and during yeast meiotic recombination (Moore *et al.*, 1999). In contrast, DNA loops containing AAG•CTT and CAA•TTG, which are not likely to be capable of hairpin formation, are readily repaired *in vivo* (Moore *et al.*, 1999). One can hypothesise that alternative DNA structures of sufficient length and threshold energy may exist long enough in the cell to allow protein recognition and binding, promoting the assembly of high-order molecular complexes, which will not only interfere with gene function, but also trap the secondary structure until repair occurs. Stable secondary structures might therefore be viewed as substrates for the DNA repair machinery, which is most certainly involved in repeat instability (Kovtun and McMurray, 2001; Manley *et al.*, 1999b; Pearson *et al.*, 1997; van Den Broek *et al.*, 2002; Chapter 8). Considered together, these observations make error-prone DNA repair, mediated by alternative structures and mismatch repair proteins, a very attractive model for triplet repeat instability. Rather than preventing recognition by repair enzymes and totally escape repair, alternative DNA structures may instead be “mis-repaired”, causing repeat expansion. In this view, slipped structures can be considered a logical candidate for mediating trinucleotide instability, and in fact there may be a number of ways by which slipped DNA can form in a mammalian cell. Any cellular process that requires transient single-stranded DNA (*e.g.* replication,

transcription, recombination or even repair) is a potential source of S-DNA structures, which may trigger a complex cascade that ends in repeat expansion. Theoretically, simple DNA breathing, could also initiate trinucleotide repeat mutation by disrupting hydrogen bonds between both strands.

Given the remarkable stability of slipped structures, one may speculate that a cellular protein is required to drive branch migration and facilitate the removal of slipped-stranded structures or interconversion between conformational isomers. We tested the ability of RecA to catalyse such a reaction, and our data reveals that this protein fails to catalyse the interconversion between structural isomers under the experimental conditions tested. Moreover, clear evidence of gel retardation was not detected. Since suboptimally paired frames, such as S-DNA structures, are known to retain higher levels of RecA than fully maximised products (Sen *et al.*, 2000), detection of band shifts might have been anticipated for low mobility DNA species, as a result of RecA binding to non-B-DNA structures containing hairpins and loop-outs. Several reasons may account for the lack of enzymatic binding and/or activity of RecA on duplex CAG•CTG trinucleotide repeats. The enzyme may be unable to interact with the (CTG)<sub>12</sub> oligonucleotide in the pre-synaptic step of the reaction, during which the protein binds and polymerises onto the single-stranded DNA prior to strand exchange. It has been reported that the folding of single-stranded DNA into secondary structure within self-hybridised sequences and base stacking causes a strong barrier for RecA binding (Bar-Ziv and Libchaber, 2001). Despite the melting step carried out prior to the reaction, the stable alternative structure adopted by short single-stranded CTG repeats (Gacy *et al.*, 1995; Mitas, 1997), may create a strong obstacle for RecA polymerisation *in vitro*. Alternatively, as observed with GT dinucleotide repeats (Dutreix, 1997), RecA may bind to single-stranded CTG sequences with great affinity, but the repetitive sequences may strongly inhibit the formation of stable synapses, which requires the alignment of homologous sequences between different molecules, and the subsequent strand exchange step. Assuming that the reaction could proceed in the absence of a single-stranded (CTG)<sub>12</sub> sequence through a four-strand exchange mechanism, previously described for a 25-nucleotide substrate (Gamper *et al.*, 2000), one may hypothesise that this mechanism may not take place with a 277 bp long double-stranded DNA target, such as that used in the assay described in Section 4.2.2.1. The duplexes may be held so tightly that the scanning required for homologous alignment might be extremely slow (Zaitsev and Kowalczykowski, 1999). Since branch migration of slipped-out trinucleotide repeats would require not only the breaking and reforming of base pairs at the junction point, but also at every base pair within the slipped-out hairpin, it may be severely impeded by intrastrand base pairs within the loop-outs. Therefore, the activation energy for breakage of long intrastrand duplex regions may be so high, that complete renaturation will be extremely time-consuming. Regardless of the reason underlying the lack of RecA-catalysed interconversion between structural isomers, it is noteworthy that an evident role for the RecA recombination system in instability was not detected on both long and short CAG•CTG repetitive tracts cloned into *E. coli* (Jaworski *et al.*, 1995; Schmidt *et al.*, 2000). More recently, it has been reported that intramolecular RecA-independent recombination

can still occur between long CAG•CTG repetitive tracts cloned into *recA*<sup>-</sup> *E. coli* strains, although at lower levels than in RecA-proficient strains (Napierala *et al.*, 2002).

Inhibition of topoisomerase type II in *Dmt*-D mouse kidney cells by novobiocin, resulted in a highly significant decrease in the levels of expansion-biased trinucleotide repeat instability in one particular cell line. Since topoisomerase enzymes affect the topology and organisation of intracellular DNA, the primary effects of their inactivation are likely to generate far-reaching ripples. Therefore it is not easy to pinpoint the specific molecular mechanism by which topoisomerase II inhibition by novobiocin affects the dynamics of trinucleotide repeats.

The fundamental need for topoisomerases in DNA metabolism derives from the double helical structure of genomic DNA. Not only does DNA structure lead to topological predicaments that must be solved by topoisomerases, but also the topological state of the DNA itself must be fine-tuned to optimise DNA function. As a consequence, these enzymes have been found to participate in nearly all cellular transactions of DNA (Wang, 1996). Topoisomerase II, in particular, unknots and untangles DNA by passing an individual intact helix through a transient double-stranded break, generated in a separate helix, thereby resolving intramolecular DNA knots as well as intermolecular tangles (Burden and Osheroff, 1998; Wang, 1996). Eukaryotic type II DNA topoisomerases were reported to interact preferentially with curved DNA and with DNA crossovers, which may possibly account for the preferential binding of eukaryotic type II enzymes to positively or negatively supercoiled DNA (Wang, 1996). Therefore it might be possible that topoisomerase II binds to trinucleotide repetitive sequences, as they accumulate greater levels of supercoiled DNA (Bacolla *et al.*, 1997; Gellibolian *et al.*, 1997), in order to resolve complex supercoiled motifs, prevent excessive supercoiling of intracellular DNA and to allow protein access. The failure to prevent excessive supercoiling may affect DNA structure and its interactions with other molecules (Nitiss, 1998). Consequently, excessive levels of supercoiling, induced by topoisomerase II inhibition, may hinder DNA recognition by proteins involved in repeat metabolism, particularly in repeat dynamics, resulting in decreased rates of expansion. In conformity with this hypothesis, topoisomerase was unambiguously identified as a component of a chromatin-remodelling factor, which uses energy to increase the general accessibility of DNA in chromatin (Varga-Weisz *et al.*, 1997). It is reasonable to imagine that the components of the mismatch repair system, may not be able to interact with inaccessible supercoiled trinucleotide repeat sequences in the absence of topoisomerase activity, which would result in decreased levels of somatic mosaicism (Manley *et al.*, 1999b; van Den Broek *et al.*, 2002; Chapter 8). Alternatively, S-DNA formation may cause superhelical tension displacement and accumulation of supercoiling with the DNA sequences flanking loop-outs and hairpin-like structures, through a topoisomerase II-dependent manner. Impaired DNA topoisomerase II activity may, therefore, inhibit S-DNA formation, and consequently repeat expansion.

The lack of a significant effect of novobiocin on the two other cell lines studied is also intriguing, and may reflect different sensitivities to the drug developed by independent cell lines. The mechanisms of natural and acquired resistance to topoisomerase inhibitors may rely on intrinsic differences between the three cell lines. The classical multidrug resistance phenotype is

associated with increased expression or activity of a transmembrane glycoprotein named Pgp, a product of the *multidrug resistant 1 (MDR1)* gene (Gottesman and Pastan, 1988). Interestingly, Pgp is highly expressed in the kidney (Fojo *et al.*, 1987), and may therefore justify the low sensitivity to novobiocin shown by D2763Kc2 and D4132K cells. In contrast, decreased *MDR1* expression/activity levels in D2763K cells might explain the greater effect of novobiocin on the dynamics of the transgenic repeat in this line. Resistance to novobiocin may also occur by any mechanism that tends to reduce the interaction between the drug and its target enzyme, such as the appearance of a mutant topoisomerase II enzyme in cultured cells (Larsen and Skladanowski, 1998), which would account for the decreased effect of novobiocin in the dynamics of trinucleotide repeats in D2763Kc2 and D4132K cells. Plus, slow growing cell populations with a low S phase fraction show increased resistance to topoisomerase inhibitors (Larsen and Skladanowski, 1998), which would explain the lack of a significant stabilising effect on the transgenic repeat carried by the slow proliferating D4132K cells, which divide twice as slow as the other cell lines studied. Alternatively, the subtle effect of novobiocin on the triplet repeat dynamics in D4132K cells may be explained by the intrinsic dynamics of the transgene in this particular cell line. This culture failed to show a great degree of repeat expansion, in contrast to the other two cell lines. In fact, at the end of the treatment, the median repeat sizes for both control and treated D4132K cells were lower than that observed for the progenitor culture. Expansion-biased repeat instability appears to be an essential requirement for a cell line to become a suitable tool in the search for genotoxic agents that may stabilise trinucleotide repetitive sequences. The intrinsic repeat dynamics of D4132K cells may therefore raise serious questions about the use of this cell line in this investigation. Nevertheless, this culture may be extremely useful in the identification of genotoxic factors that are able to destabilise trinucleotide repeats and induce somatic mosaicism. Finally, DNA repair activity levels may also be associated with resistance to topoisomerase inhibitors. Resistant cell lines developed *in vitro* are more likely to have attenuated checkpoints due to p53 mutations (Ogretmen and Safa, 1997). Likewise, mutations in the mismatch repair genes result in the development of resistance to topoisomerase inhibitors (Fedier *et al.*, 2001). Marked culture-to-culture heterogeneity in mismatch repair activities may result in different levels of sensitivity towards topoisomerase inhibitors displayed by distinct cells and therefore explain the differing results obtained with the three cell lines studied here. For that reason one must be aware that rather than a direct consequence of topoisomerase II inhibition, the decreased levels of expansion-biased repeat instability might have resulted from a selection for mutations in the DNA MMR genes. The accumulation of MMR-deficient cells in culture would consequently lead to a decrease in repeat instability, in agreement with the recently reported functions of mismatch repair genes as enhancers of somatic mosaicism in mouse cells (Manley *et al.*, 1999b; van Den Broek *et al.*, 2002; Chapter 8). Nevertheless, the effect of novobiocin on triplet repeat dynamics should be extended to additional cell lines, and possibly tested *in vivo*, to clarify the involvement of topoisomerase II in trinucleotide repeat metabolism,

Although novobiocin was previously reported to inhibit replicative DNA synthesis in some human and rodent cell lines at concentration higher than those used in this study (Mattern and

Scudiero, 1981), some contradictory reports claimed that inactivation of DNA topoisomerase II does not prevent DNA synthesis (Andoh and Ishida, 1998). It is important to note that the population doubling time of treated cultures did not increase significantly, suggesting that the DNA replication in the cells lines used in this study is not sensitive to the amount of novobiocin added to the growth medium. Therefore, it can be concluded that the novobiocin-induced decrease in repeat expansion is not mediated by a decrease in the rate of replicative DNA synthesis.

If alternative DNA structures were the ultimate molecular bases of trinucleotide repeat disorders, altering such structural dynamics could influence the mutation dynamics of these sequences and serve as a potential therapeutic strategy, and topoisomerase inhibitors may represent a good chemical candidate. The results obtained with novobiocin not only corroborate this idea, but also bring topoisomerases as new players into the field of trinucleotide repeat dynamics. There is also evidence that some topoisomerases might interact with one or more of the mammalian helicases, such as Bloom (BLM) and Werner (WRN) helicases (Champoux, 2001; Duguet, 1997). It is noteworthy that WRN helicase has also been implicated in trinucleotide repeat metabolism, as it efficiently unwinds single-strand CGG tetraplex structures (Fry and Loeb, 1999), and through a concerted interaction with DNA polymerases is able to resolve tetraplex secondary structures in DNA templates and restore full-length DNA synthesis of CGG-containing sequences (Kamath-Loeb *et al.*, 2001). In addition, tetraplex structures of CGG oligomers interspersed by AGG interruptions are unwound by human WRN helicase at higher rates and to a greater extent than tetraplexes consisting of monotonous uninterrupted CGG•CCG repeats (Weisman-Shomer *et al.*, 2000), suggesting a possible molecular mechanism by which AGG•CCT interruptions may stabilise CGG•CCG tracts and restrict their expansion. The significance of these protein interactions is still poorly understood, but they support the involvement of a multi-component ensemble in triplet repeat biology.

The notion that secondary structure mediates repeat expansion in double-stranded DNA may finally be more of a fact than a hypothesis. A greater understanding of both the triplet repeat DNA structures and the molecular mechanisms for spontaneous mutation will be required before we can understand which of the molecular mechanisms currently under discussion are responsible for trinucleotide repeat instability and development of human disease.

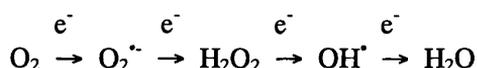
## 5. Genotoxic effects of oxidative stress on trinucleotide repeat dynamics

### 5.1. Introduction

Oxidative stress, a phenomenon resulting from the imbalance between oxygen species and protection against oxidants, contributes to spontaneous mutagenesis in somatic cells (Rossman and Goncharova, 1998). Endogenous oxidative damage to DNA is extensive and has been estimated to account for 10,000 hits per cell per day in humans (Ames *et al.*, 1993), and has been implicated in the aetiology of cancer and ageing, as well as several other human diseases, particularly neurodegenerative conditions (Ames *et al.*, 1993; Beckman and Ames, 1997; Croteau and Bohr, 1997; Gracy *et al.*, 1999).

Many of the biological effects of oxidative damage are mediated by the highly reactive oxygen species (ROS). Metabolism, like other aspects of life, involves trade-offs. Oxidant by-products of normal metabolism cause extensive cellular damage. There are various intra- and extracellular sources of oxygen radicals, the major intracellular sources probably being the leakage associated with the reduction of oxygen to water during mitochondrial respiration and the by-products of peroxisomal metabolism (Ames *et al.*, 1993; Burcham, 1999; Raha and Robinson, 2000).

The unpaired orbitals of oxygen can sequentially accommodate single electrons, to yield superoxide radical ( $O_2^{\cdot-}$ ), hydrogen peroxide ( $H_2O_2$ ), the extremely reactive hydroxyl radical ( $OH^{\cdot}$ ) and finally water ( $H_2O$ ) (Davies, 1999):



The leakage of partially reduced oxygen molecules from the mitochondrial electron chain is around 2%, yielding about  $2 \times 10^{10}$  superoxide and hydrogen peroxide molecules per cell per day (Ames *et al.*, 1993; Rossman and Goncharova, 1998). Free superoxide is relatively unreactive. However, superoxide undergoes dismutation (either spontaneously or via enzyme-catalysed reactions) to produce hydrogen peroxide. Hydrogen peroxide is not itself an oxygen radical and is also relatively stable. Nevertheless, hydrogen peroxide is a diffusible, latent oxygen species that can move great distances within cells to react with transition metals, such as iron, to produce extremely reactive hydroxyl radicals by the Fenton reaction (Henle and Linn, 1997; Raha and Robinson, 2000):



Hydroxyl radicals are not only capable of damaging all macromolecules in the cell, including phospholipids, proteins and DNA, but also of causing DNA-protein crosslinks, as well as

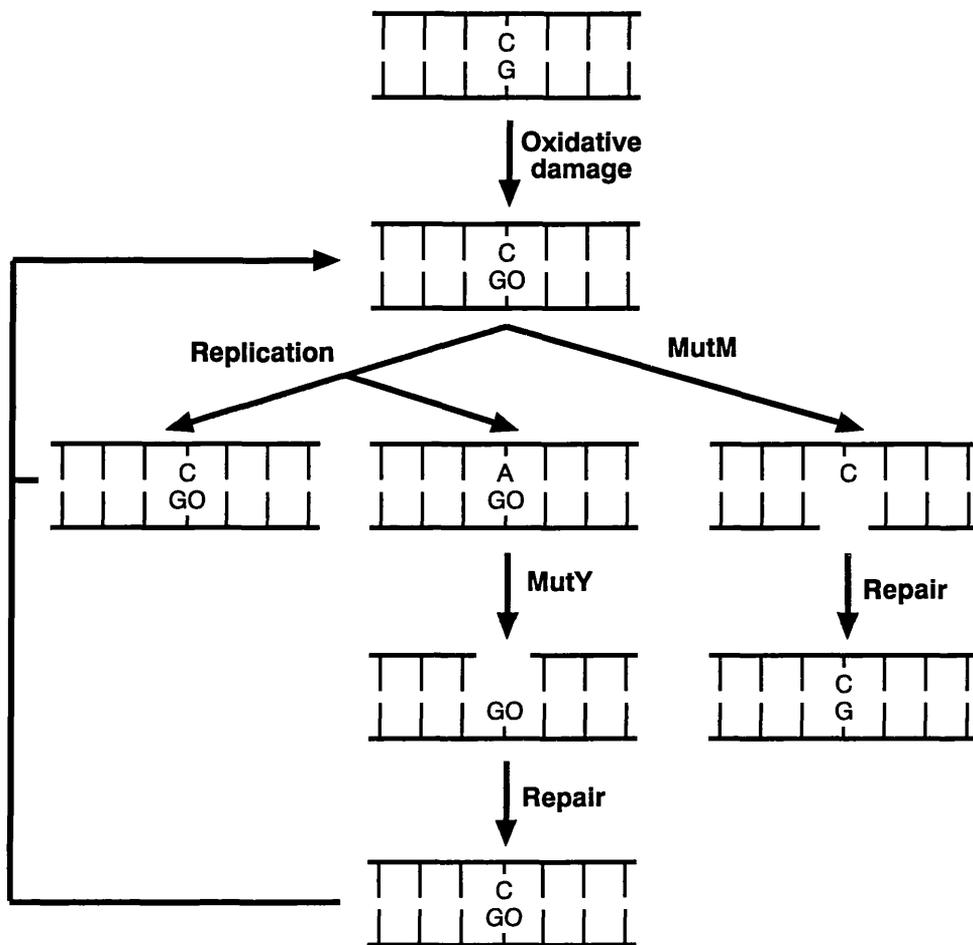
single and double strand DNA breaks (Burcham, 1999; Epe, 1996; Henle and Linn, 1997; Marnett, 2000). Much endogenous DNA damage arises from the direct attack of intermediates of oxygen reduction on the bases or the deoxyribose backbone of DNA. Alternatively, oxygen radicals can oxidise other cellular components such as lipids, sugars, estrogens, amino acids, heme groups, to generate reactive intermediates capable of attacking DNA bases, giving rise to bulky adducts, which distort the B-DNA conformation (Burcham, 1999; Marnett, 2000; Moller and Wallin, 1998). Radical attack on the bases results primarily in the addition of a hydroxyl group to the electron-rich double bonds, resulting in more than 30 different types of base alterations, with some authors reporting more than 100 modifications (Henle and Linn, 1997; Marnett, 2000). The mutagenic consequences of only a few of these adducts are known. The deoxyguanosine oxidation product 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxoG), for instance, can be produced in DNA by hydroxyl radical attack, and if not repaired can cause G•C→T•A transversions. Another highly mutagenic oxidised base, 5'-hydroxy-2'-deoxycytidine, causes C•G→T•A transitions, while thymine glycol causes T•A→C•G transitions (Marnett, 2000).

In spite of the constant presence of ROS in the cell and the generation of a diverse range of premutagenic DNA adducts by oxygen metabolism, the background level of oxidative damage in normal mammalian cells is probably very low, since antioxidant defences and DNA repair pathways have evolved to maintain DNA integrity. Minimal damage accumulation is achieved by multiple interacting systems of antioxidant compounds, antioxidant enzymes, damage removal enzymes and repair enzymes (Collins and Horvathova, 2001). Many defence mechanisms have, therefore, evolved to limit the levels of reactive oxidants and the damage they inflict. The first level of cellular responses to oxidative stress consists of a series of antioxidant scavengers and enzymes that rapidly remove reactive oxygen species, thereby avoiding cellular damage. Small dietary antioxidants such as vitamin C (ascorbate), vitamin E (tocopherol) and carotenoids have antioxidant activities. Antioxidant compounds are sacrificed to oxidation in order to directly protect more important cellular components. Antioxidant enzymes such as superoxide dismutases, glutathione peroxidases, and quinone reductases act catalytically to convert oxidants to less reactive species. Superoxide dismutase reduces superoxide anion radicals to molecular oxygen and hydrogen peroxide. Catalase and glutathione peroxidase reduce hydrogen peroxide or other hydroperoxides to water or corresponding hydroxyl compounds (Ames *et al.*, 1993; Davies, 1999; Henle and Linn, 1997).

The production of endogenous oxidants is expected to increase under the conditions of oxidative stress, and since the first line of antioxidant defences developed by animals is not perfect, some DNA is consequently oxidised. Oxidatively damaged DNA is repaired by enzymes that excise the lesions. Repair of oxidative damage becomes especially important, given that endogenous oxidatively damaged bases are continuously generated in the DNA of normal, healthy individuals, being influenced by numerous environmental factors. Once DNA nucleoside damage is manifested, the lesion must be recognised, removed and replaced with normal nucleotides (Ames *et al.*, 1993; Beckman and Ames, 1997). Due to the high levels of endogenous oxidative damage, mammalian cells evolved multiple repair mechanisms to survive the daily insults. Most DNA

lesions are repaired by either of the two excision repair processes: the base excision repair (BER) or the nucleotide excision repair (NER). The two processes are similar in the overall mechanism: a patch of DNA containing the lesion is removed, a DNA polymerase synthesises a new strand using the complementary strand as template, and a DNA ligase seals the last nick. However these two repair pathways utilise different enzymes and differ in detail (Moller and Wallin, 1998). BER is initiated by DNA glycosylases, a class of enzymes that recognise a specific set of modified bases, such as 8-oxoG or thymine glycol. Glycosylases cleave the *N*-glycosylic bond between the modified base and the sugar moiety, generating an abasic site (AP site). Following the glycosylase step, AP endonucleases are required to remove the 3'-deoxyribose moiety and generate a 3'-hydroxyl group, which can be extended by a DNA polymerase. The process is completed by a DNA ligase, which rejoins the free DNA ends (Croteau and Bohr, 1997; Henle and Linn, 1997; Krokan *et al.*, 1997; McCullough *et al.*, 1999). NER is the most complicated of the excision repair systems. The system acts upon a wide range of alterations that result in large distortions in DNA and removes the lesion as part of an oligonucleotide. A new DNA sequence is synthesised using the intact strand as template (Cleaver *et al.*, 2001; Henle and Linn, 1997). Many studies show that NER plays a fundamental role in repair of oxidative lesions that are not substrates for BER. In addition, NER may also act as a backup system, if BER happens to be saturated (Moller and Wallin, 1998).

One of the most frequent mutagenic base lesions, induced by oxygen free radicals, is 8-oxoG. This altered base can mispair with adenine as well as cytosine residues, leading to a greatly increased frequency of spontaneous G•C→T•A transversions mutation (Burcham, 1999; Marnett, 2000). In bacterial cells, three enzymes act to prevent spontaneous mutagenesis induced by 8-oxoG (Figure 5.1). MutT (a 8-oxo-dGTPase) provides the first line of defence by eliminating 8-oxo-dGTP from the dNTP pool. If 8-oxo-dGTP escapes MutT activity, this oxidised nucleotide triphosphate may be incorporated opposite either adenine or cytosine. Here, MutM constitutes a second line of defence by removing 8-oxoG incorporated opposite cytosine or formed by the oxidation of DNA guanine. MutM repairs 8-oxoG incorporated opposite adenine with low efficiency. If both of these defence levels are bypassed, and the DNA is replicated, MutY provides a third level of defence by removing adenine incorporated opposite 8-oxoG (Burcham, 1999; Krokan *et al.*, 1997; Marnett, 2000; McCullough *et al.*, 1999; Michaels and Miller, 1992). The human and the mouse 8-oxoguanine *MutM* glycosylase gene (human *8-oxoG DNA glycosylase*, *OGG1*, or mouse *Ogg1*) have been cloned, and they are ubiquitously expressed in a variety of organs and exhibit a repair activity that selectively excises 8-oxoG from oxidatively damaged DNA (Aburatani *et al.*, 1997; Arai *et al.*, 1997; Bjoras *et al.*, 1997; Lu *et al.*, 1997; Radicella *et al.*, 1997; Roldan-Arjona *et al.*, 1997; Rosenquist *et al.*, 1997). Homozygous *ogg1*<sup>-/-</sup> null mice are viable and exhibit a moderately, but significantly, elevated spontaneous mutation rate especially in slowly proliferative tissues (Klungland *et al.*, 1999). Human *mutY* and *mutT* homologues, *MYH* and *MTH*, respectively, have also been identified (McGoldrick *et al.*, 1995; Sakumi *et al.*, 1993; Slupska *et al.*, 1996). The cloning of the human homologues has strongly suggested that human cells possess repair systems, similar to those of bacteria and yeast, to protect the genome from the mutagenic



**Figure 5.1. The bacterial 8-oxoG base excision repair system.**

Oxidative damage can lead to 8-oxoG (GO) lesions in DNA. The 8-oxoG lesions can be removed by MutM protein, and subsequent repair can restore the original G•C basepair. If the 8-oxoG lesion is not removed before replication, synthesis by replicative DNA polymerases is frequently inaccurate, leading to misincorporation of A opposite the 8-oxoG lesion. MutY removes the misincorporated adenine from the 8-oxoG•A mispairs. Repair polymerases are much less error prone during synthesis and they usually generate a 8-oxoG•C mismatch, a substrate for MutM. A third repair enzyme, MutT, is active on 8-oxodGTP and hydrolysis it to 8-oxodGMP, effectively removing the oxidised triphosphate from the deoxynucleotide pool, otherwise, inaccurate replication could result in the misincorporation of 8-oxodGTP opposite template A residues, leading to 8-oxoG•A mispairs.

lesions induced by endogenous and environmental ROS.

However, a greater level of complexity is present in higher eukaryotes by the introduction of a mitochondrial genome. Mitochondrial genomes encode ribosomal mitochondrial RNA, all the tRNAs and 13 polypeptides, all of which are subunits of the enzymatic complexes involved in oxidative phosphorylation. It has been shown that oxidative damage to mitochondrial DNA occurs more frequently than damage to nuclear DNA (Yakes and Van Houten, 1997). The increased susceptibility of the mitochondrial genome to oxidative damage could be due to a lack of repair enzymes, a lack of histones protecting mitochondrial DNA (Enright *et al.*, 1992), and the proximity of mitochondrial DNA to oxidants generated during oxidative phosphorylation (Ames *et al.*, 1993; Yakes and Van Houten, 1997). Any impairment of the mitochondrial genome will cause dysfunction of the respiratory chain and hence reduce ATP synthesis (Raha and Robinson, 2000; Yakes and Van Houten, 1997). Given the essential mitochondrial function of generating energy, the integrity of mitochondrial DNA must be maintained. Different forms of OGG1 and MYH enzymes, generated by alternative splicing, have been found in the mitochondria and nuclei (Nishioka *et al.*, 1999; Takao *et al.*, 1998; Takao *et al.*, 1999), strongly suggesting that the same repair system against 8-oxoG mutagenic lesions operates in mitochondria. In addition, mitochondria also contain an error avoidance *mutT* homologue (Kang *et al.*, 1995a). However, NER as it exists in the nucleus does not operate in mitochondria. It is therefore apparent that mitochondria possess the capability to remove oxidative DNA damage, whether they use the same nuclear repair proteins or different ones (Croteau and Bohr, 1997).

Given the high levels of endogenous oxidative damage, both BER and NER systems may be overwhelmed. Therefore, the MMR pathway has also been implicated in the repair of DNA oxidative damage, by correcting newly acquired ROS-induced mutations after DNA replication, immediately before mitosis (Gasche *et al.*, 2001; Lin *et al.*, 2000b). Indeed, mutations in bacterial MMR genes result in increased microsatellite instability of repetitive sequences upon oxidative damage, implicating the involvement of the MMR repair system in mutation avoidance induced by oxidative stress (Jackson and Loeb, 2000). It has been speculated that being an abundant source of endogenous mutagens, ROS have a major impact on genetic integrity, particularly promoting instability of microsatellite sequences. Several pieces of evidence suggest that DNA oxidative damage by hydrogen peroxide is sufficient to promote instability of microsatellite repetitive sequences, even in the presence of proficient DNA repair systems. Exposure of *E. coli* to low levels of hydrogen peroxide increases the frequency of expansions and deletions within dinucleotide repetitive sequences (Jackson and Loeb, 2000). Increased rates of frameshift mutations, exclusively localised within microsatellite sequences, have also been reported following direct exposure of plasmid DNA to hydrogen peroxide *in vitro*, and subsequent transformation into repair-proficient *E. coli* (Jackson *et al.*, 1998). This observation confirms that the increased mutation rate is a direct consequence of the DNA damage induced by hydrogen peroxide, rather than a secondary effect, mediated by an intermediate pathway. Similarly, repeated exposure of mammalian cells to hydrogen peroxide leads to a 9-fold increase in the mutations rate of CA•GT dinucleotide microsatellite sequences, cloned into plasmids and transfected into human colorectal cancer cells

(Gasche *et al.*, 2001). In addition, the CA•TG microsatellite mutation rate is 27 times higher than the spontaneous mutation frequencies, following the exposure of lung cancer cell lines to hydrogen peroxide (Zienolddiny *et al.*, 2000). In the same way as bacteria, MMR-proficient mammalian cells are less susceptible to hydrogen peroxide-induced mutations than MMR-deficient cells, indicating that the mammalian MMR enzymes play a crucial role in the repair of mutations after oxidative stress (Gasche *et al.*, 2001). The mechanism by which oxidative stress-induced DNA damage promotes microsatellite instability has not yet been identified. One mechanism could be mediated by pausing of DNA polymerase at sites of oxidative base damage (Feig and Loeb, 1993), while another mechanism could be dependent on single strand breaks. Both provide the opportunity for strand displacement, formation of slipped-stranded intermediates at different positions, generating insertion or deletion loops (Jackson *et al.*, 1998). Furthermore it has been suggested that oxidative stress may alter MMR function and allow mutations to accumulate over time. Alternatively, the higher levels of DNA damage during oxidative stress may also overwhelm the MMR system allowing replication errors to pass the G<sub>2</sub> cell cycle checkpoint (Gasche *et al.*, 2001).

Given the effect of hydrogen peroxide-induced oxidative stress on dinucleotide microsatellite instability, one may speculate that oxidative stress might also modify the dynamics of trinucleotide repeats in mammalian cells. Previous independent reports, when considered together, appear to point to an association between levels of oxidative stress and repeat dynamics in mouse models of trinucleotide repeat instability. Increased levels of 8-oxoG and lipid peroxidation were found in the striatum of HD knock-in R6/1 and R6/2 transgenic mice, relative to their wild-type littermates (Bogdanov *et al.*, 2001; Perez-Severiano *et al.*, 2000). No significant differences in lipid peroxidation were found between cerebella of R6/1 transgenic and wild-type mice at any age (Perez-Severiano *et al.*, 2000). Interestingly, the striata of both R6/1 and R6/2 mice exhibit the greatest levels of trinucleotide repeat instability, in particular, higher degrees of somatic mosaicism have been reported in the striatum, relative to the cerebellum of these mice (Mangiarini *et al.*, 1997). Taken together, these data may indicate that increased levels of oxidative stress parallel greater degrees of trinucleotide repeat instability. In addition, the *Dmt-D* mouse tissues that exhibit the highest levels of somatic mosaicism (kidney, liver and brain (Fortune *et al.*, 2000)) also reveal the highest levels of antioxidant defences. When mouse tissues are ranked from highest to lowest regarding the activity of superoxide dismutase, catalase and glutathione peroxidase enzymes, liver and kidney exhibit the highest average rank, followed by brain (Grankvist *et al.*, 1981; Ibrahim *et al.*, 2000). Moreover, kidney also reveals great expression levels of extra cellular superoxide dismutase (Ookawara *et al.*, 1998). Similarly, the levels of non-enzymatic antioxidants, such as glutathione, vitamin C and vitamin E are higher in kidney and liver, compared to other mouse tissues (Ibrahim *et al.*, 2000). These data suggest that these particular tissues accumulate protective enzymes and antioxidants, possibly as a result of their own high levels of endogenous oxidative stress. Yet, despite the greatest antioxidant defences in kidney, this tissue still exhibits the high levels of malondialdehyde (Ibrahim *et al.*, 2000), a lipid oxidation product able to attack DNA with major implications on DNA integrity (Burcham, 1999; Marnett, 2000; Moller and Wallin, 1998). In

summary, a correlation between the redox status of a tissue and trinucleotide repeat dynamics appears to emerge from the analysis of previously published data.

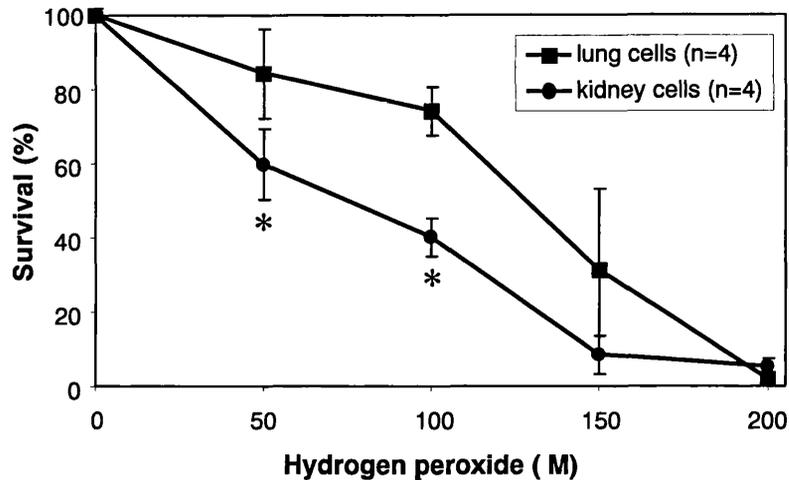
Oxidative damage might therefore be a key factor in modifying the mutation rate of trinucleotide repetitive sequences, possibly contributing to the age-dependent and tissue-specific mutation profile, characteristic of trinucleotide repeat disorders. In order to test the hypothesis that ROS may create transient hypermutable intermediates, which may affect the faithful maintenance of triplet repeats and lead to instability, *Dmt-D* cell cultures were exposed to chemical compounds known to increase the intracellular levels of oxidative stress, and following the treatment, trinucleotide repeat size variability was monitored by sensitive SP-PCR techniques.

## 5.2. Results

### 5.2.1. *Association between levels of somatic instability and sensitivity to oxidative stress in culture*

In order to test for a possible association between the intrinsic levels of tissue-specific trinucleotide repeat instability, exhibited by a cell line, and the sensitivity of the culture to oxidative damage, two different cell lines were treated with increasing concentrations of hydrogen peroxide. D2763 lung and kidney cells (D2763L and D2763K cultures, respectively, Table 3.2) were selected to perform this study, since they were derived from the same mouse, and exhibited distinct levels of somatic instability: whereas the transgene remains remarkably stable in cultured D2763L lung cells, great levels of expansion-biased repeat instability were detected in D2763K kidney cells (Section 3.2.4). Replicates of each culture were treated with concentrations of hydrogen peroxide ranging from 50 to 200  $\mu\text{M}$ , for 8 hours, at 37°C and 5% (v/v)  $\text{CO}_2$ . Following the incubation period, cell viability was assessed according to the acridine orange and ethidium bromide method (Section 2.4.6), to determine the percentage of cell survival following each treatment (Figure 5.2). As expected, an increase in cell death was measured with increasing levels of hydrogen peroxide added to the culture, independently of the cell line. More interestingly, the graph reveals a marked difference in the sensitivity of the two cell lines to the exogenous oxidant. The resistance of lung cells to 50 and 100  $\mu\text{M}$  hydrogen peroxide proved significantly higher ( $p < 0.05$ , two-tailed *t*-test), with levels of cell viability ~1.5- and ~2-fold higher than kidney cultures, respectively.

This observation suggests a parallel between the levels of sensitivity to hydrogen peroxide-induced oxidative stress, and the intrinsic levels of trinucleotide repeat instability exhibited by the different cell lines under standard growth conditions, indicating a possible role for ROS in triplet repeat metabolism.



**Figure 5.2. Sensitivity of *Dmt-D* cultured cells to hydrogen peroxide.**

Four independent *Dmt-D* lung and kidney cell cultures, derived from the same mouse, were treated for eight hours with increasing concentrations of hydrogen peroxide, ranging from 50 to 200  $\mu\text{M}$ . At the end of the treatment, cell viability was measured by the acridine orange and ethidium bromide method, and plotted as a function of the concentration of hydrogen peroxide in the culture medium. The graph shows a decrease in cell survival with increasing concentrations of hydrogen peroxide. Kidney cells exhibited greater sensitivity to hydrogen peroxide than lung cell cultures. The difference between the sensitivity of the two cell lines to hydrogen peroxide proved statistically significant at 50 and 100  $\mu\text{M}$  (\*,  $p < 0.05$ , two-tailed *t*-test).

### **5.2.2. Induction of trinucleotide repeat size variability in vitro by hydrogen peroxide: a preliminary study**

To further test the involvement of ROS in the molecular mechanisms driving trinucleotide repeat dynamics, a preliminary study was performed on three *Dmt-D* cell lines. Two replicate cultures were established from the same progenitor cell population: one was treated with 100  $\mu\text{M}$  hydrogen peroxide (unless otherwise stated) for up to 58 population doublings, while the other was cultured under standard growth conditions, in the absence of the oxidative chemical. At the end of the treatment, the repeat size variability within each culture was determined by sensitive SP-PCR procedures. Since the hydrogen peroxide treatment affected the proliferative capacity of the cells, and resulted in a marked increase in the population doubling time, the repeat size variation in treated cells was not only compared with controls grown for the same number of population doublings, but also with controls maintained in culture for the same period of time.

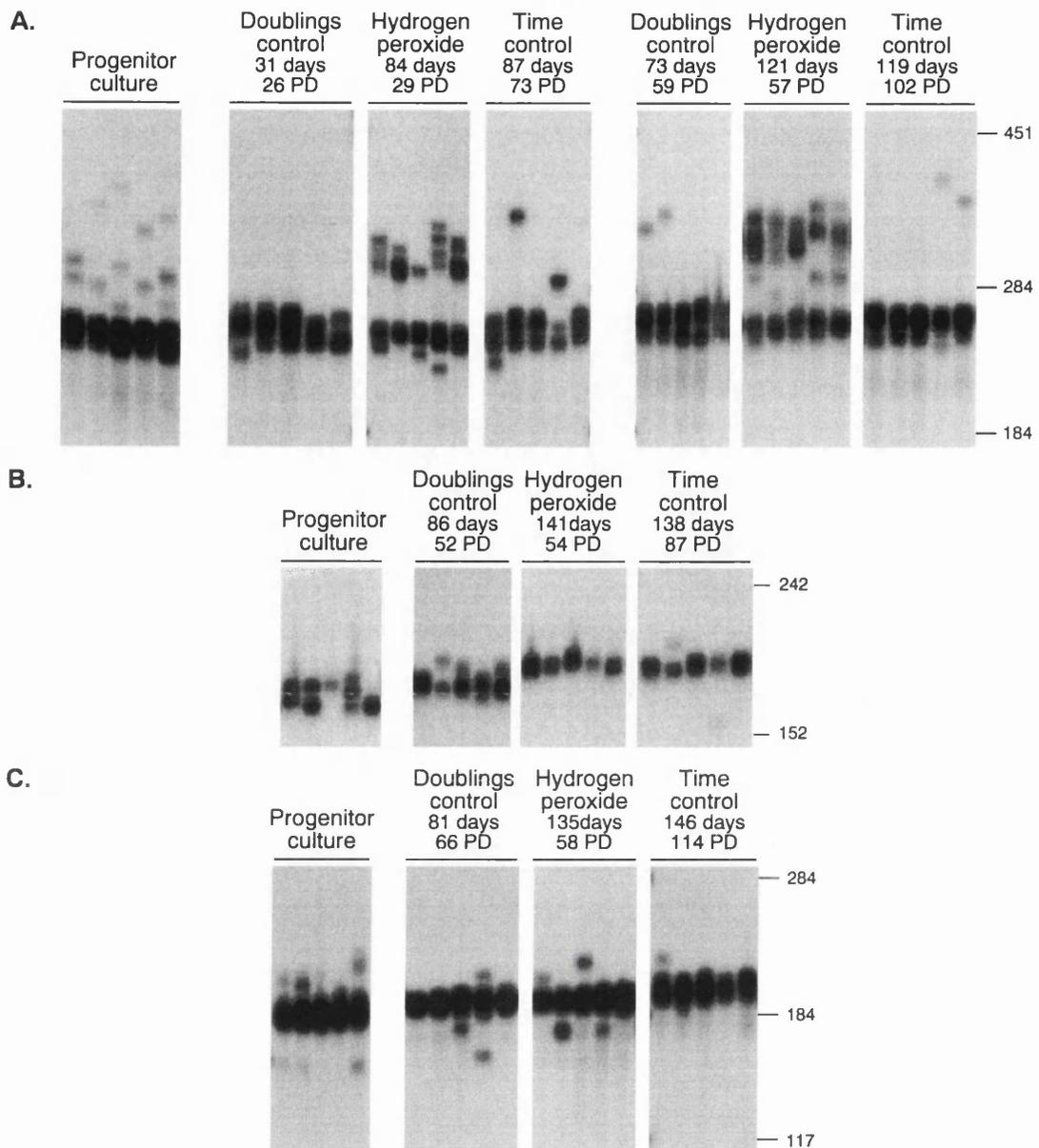
It should be stressed at this point that an extensive single molecule analysis was not performed at this initial stage, as the aim of this study was the gathering of experimental qualitative evidence indicative of a possible role of oxidative stress in the metabolism of triplet repeats. Thus,

no statistical tests were performed, and the average allele sizes presented were estimated based on the sizing of a low number of transgene molecules, usually between ten and 20.

#### 5.2.2.1. D2763K kidney cell line

The progenitor kidney cell culture derived from a six-month-old *Dmt-D* mouse had previously been analysed and reported to exhibit great levels of expansion-biased trinucleotide repeat instability (Section 3.2.4). Hydrogen peroxide treatment resulted in reduced cell proliferation with an increase of 80% in the population doubling time.

When the hydrogen peroxide treatment was established, the progenitor culture, which had previously been growing for 215 days and 91 population doublings, consisted of a major subpopulation of cells carrying an average allele size of ~260 repeats. Less than ~10% of the cells carried longer alleles, containing up to 400 repeats (Figure 5.3.A). A lower level of repeat size variability was retained by the controls throughout this study, indicating that the repeat length variation was primarily being shaped by a major selective sweep, as discussed previously in Chapter 3. In contrast to the control cells, treatment of D2763K kidney cultures with 100  $\mu$ M hydrogen peroxide for 121 days and 57 population doublings cultures, resulted in a striking increase in the overall median repeat length, as a result of the expansion of a second subpopulation of cells within the culture. This population comprised ~50-60% of the total number of cells, carrying longer alleles with an average size ~330 repeats, estimated by the sizing of ten to 20 individual molecules. Although a small subset of large expansions was present in both controls, these mutants occurred at much lower frequency relative to hydrogen peroxide exposed cells (Figure 5.3.A), even though time control cultures had undergone nearly twice as many population doublings as the treated culture. To test for the progressive effect of hydrogen peroxide treatment on repeat dynamics and the continuous accumulation of longer alleles under conditions of oxidative stress, repeat size variability was also assessed at an intermediate time point, following 84 days and 29 population doublings of treatment. By this stage an emergent second population of larger alleles was already noticeable in treated cultures, nevertheless it consisted of only ~30-40% of the total number of cells. An accurate comparison between the average repeat sizes of the longer alleles at the two time points could not be established, since extensive single molecule sizing was not performed at this stage. Nevertheless, an average repeat size of ~320 repeats was estimated based on the sizing of ten to 20 individual molecules collected at this intermediate point. This very preliminary bulk SP-PCR analysis not only suggests a significant growth of the population of cells carrying longer alleles, but also a steady increase in the repeat number carried by those cells over the second half of the treatment, indicating that oxidative stress induced by hydrogen peroxide might strongly affect the trinucleotide repeat size profile in culture in a progressive time-dependent way.



**Figure 5.3. Preliminary analysis of CAG-CTG repeat instability in *Dmt-D* kidney and eye cells treated with hydrogen peroxide.**

Independent *Dmt-D* cells lines, showing different levels of trinucleotide repeat instability, were treated with 100  $\mu\text{M}$  (D2763K and D3111K kidney cell lines) or 200  $\mu\text{M}$  (D2763E eye cell line) hydrogen peroxide for 54 to 58 population doublings (PD) and variable periods of time. SP-PCR analyses were performed to assess the repeat size variability in treated cells, in control cultures maintained under standard conditions for a similar number of population doublings (doublings control) or days in culture (time control), and in the progenitor culture, from which all cultures were derived at day zero. The autoradiographs show representative SP-PCR amplifications of DNA extracted from each cell sample. (A) A dramatic accumulation of large expansions was observed in *Dmt-D* kidney cells (D2763K), which had previously shown higher levels of somatic mosaicism under standard growth conditions. Also note the progressive increase proportion of longer alleles from 84 to 121 days of treatment. (B) The treatment of a second kidney cell line (D3111K), exhibiting modest levels of triplet repeat instability, resulted in an apparent increase in the average repeat sizes, when compared to the doublings controls. (C) Trinucleotide repeat instability in eye cells (D2763E) was not greatly affected by the treatment. An average of five to 20 transgene molecules were amplified in independent reactions. The molecular weight markers are shown on the right, after conversion into repeat number.

### 5.2.2.2. D3111K kidney cell line

The ability of hydrogen peroxide to induce increased levels of repeat size variability was studied in another kidney cell line, which exhibits moderate levels of repeat instability under standard growth conditions (D3111K cell line, Table 3.2 and Section 3.2.5). The progenitor culture had been growing for 166 days and 40 population doublings prior to the establishment of two replicates and the beginning of the treatment, as described above. At that time point D3111K cells showed a bimodal repeat size distribution with two peaks of variability with average repeat sizes of ~160 and ~175 repeats (Figure 5.3.B). Following hydrogen peroxide treatment for 141 days and 54 population doublings, which resulted in a 25% increase in population doubling time, D3111K treated cells showed low levels of repeat size variability, with no signs of multiple peaks of repeat length variation. However, the average repeat size had increased to ~200 units (Figure 5.3.B), in contrast with the shorter alleles detected in both doublings (average repeat number of ~180) and time controls (average repeat number of ~195 repeats). Given that extensive single molecule sizing was not performed, one cannot comment on the statistical significance of these observations. Regardless of the lack of quantitative data, the results appear to reveal a subtle effect of hydrogen peroxide treatment on the repeat dynamics in cultured D3111K kidney cells.

### 5.2.2.3. D2763E eye cell line

Finally, a similar study was performed with a *Dmt-D* eye cell line (D2763E line, Table 3.2), which exhibited the lowest levels of trinucleotide repeat instability among the three lines selected to carry out this study. Assuming a lower sensitivity to hydrogen peroxide, given the association between intrinsic levels of trinucleotide repeat instability and resistance to the oxidant chemical (Section 5.2.1), D2763E eye cells were expected to survive under higher levels oxidative stress. Consequently, and also to test a possible dose-response association between the levels of exogenous hydrogen peroxide and the effect on the repeat dynamics, this cell line was treated with 200  $\mu\text{M}$  of hydrogen peroxide, resulting in an increase of 86% in the population doubling time. The progenitor culture carried an average repeat number of ~185 units, and had been growing for 196 days and 122 population doublings *in vitro*. Exposure to hydrogen peroxide for 135 days and 58 population doublings, did not result in any evident changes in the trinucleotide repeat profile detected by high DNA input SP-PCR analysis (Figure 5.3.C). The average repeat size increased by 5 units up to ~190 repeats, similarly to the increase exhibited by the doublings control cells. However, time controls, whose cells underwent twice as many population doublings as the treated cells, had gained 10 repeats, displaying an average allele length of ~195 repeats. Differences in the repeat size heterogeneity within each culture could only been established by single molecule analysis, which was not performed at this stage. In summary, the results obtained with D2763E eye cells failed to reveal a detectable association between oxidative stress and altered trinucleotide repeat dynamics.

Considered together, the data collected from the three cell lines may suggest that high levels of oxidative stress, induced by the addition of chemicals to the cell growth medium, may

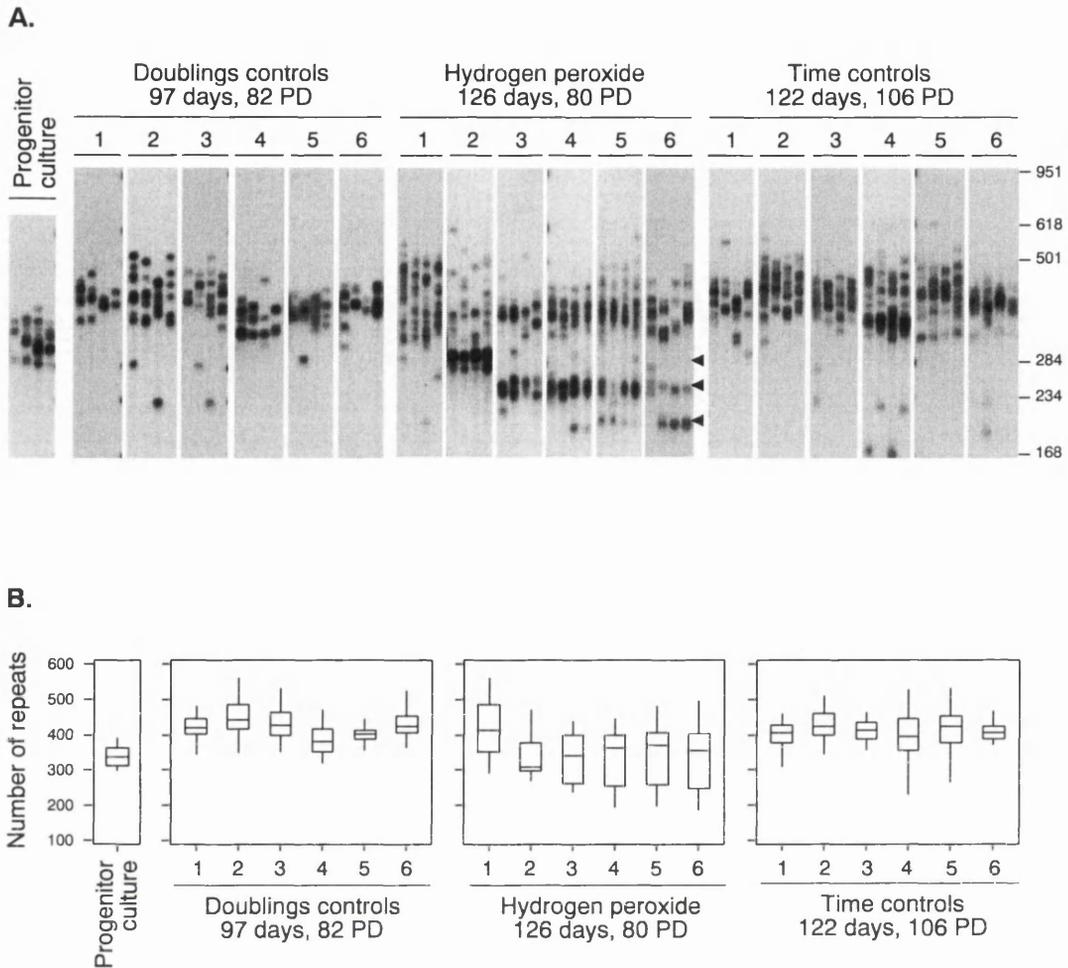
mediate triplet repeat metabolism, and in particular affect levels of somatic mosaicism detected in culture. Furthermore, the very preliminary analysis performed on D2763E eye cells, supports the view that a cell line must exhibit a measurable degree of intrinsic trinucleotide repeat instability under standard growth conditions, to increase the likelihood that a visible effect (if any) of an environmental agent upon the stability of trinucleotide repeats is detected.

### **5.2.3. Quantitative analysis of the effect of hydrogen peroxide-induced oxidative stress on the dynamics of expanded CAG•CTG repeats**

#### **5.2.3.1. D2763Kc2 kidney cell line**

Encouraged by the preliminary results described in Section 5.2.2, we sought to quantify the effect of oxidative stress induced by hydrogen peroxide on the stability of trinucleotide repeats in cultured mouse cells. To this end D2763Kc2 kidney cells were selected to carry out a detailed analysis, since they presented two major advantages relative to other cell lines. First, being a clonal cell line established from a single cell, differences in the genetic background between individual cells are minimised and expected to have a minor influence on the outcome of the treatment. Second, the transgenic repeats have been shown to expand rapidly in this cell line under standard conditions, being more likely to be affected by the addition of hydrogen peroxide to the growth medium, in line with the preliminary observations previously presented. To avoid a major effect of population fluctuations (selective sweeps, Chapter 3) and the consequent artefacts they may generate, six parallel replicate cultures were established from a single progenitor population, and treated with 100  $\mu$ M hydrogen peroxide for 126 days, corresponding to 80 population doublings. As a result of the treatment, an increase of 36% in the population doubling time was observed. Despite the previous observation (Chapter 3) that different expansion rates do not necessarily rely on differences in the proliferation capacity between cell lines, two sets of control cultures were established as usual: time and doublings controls. Furthermore, the expansion rates corrected for both time and population doublings, were compared to the closest control in terms of days in culture, or population doublings, respectively.

At the end of the treatment the repeat size variability within each culture was assessed by SP-PCR (Figure 5.4.A). Despite great variation between treated replicate cultures, the amplification of an average of five to 30 transgene molecules per reaction illustrated a marked decrease in the average repeat size in treated cells, when compared with both controls, in contrast with the preliminary results obtained with both D2763K and D3111K kidney lines. Single molecule analysis allowed the quantification of this effect. Twenty to 80 individual molecules were individually sized and the repeat lengths confirmed that hydrogen peroxide treatment resulted in a decreased rate of expansion, along with a dramatic increase in the frequency of cells containing large deletions (Figure 5.4.B). The median rate of expansion of 0.180 repeats per day, in hydrogen peroxide treated cells, was significantly lower than the median gain of 0.615 repeats per day in the controls



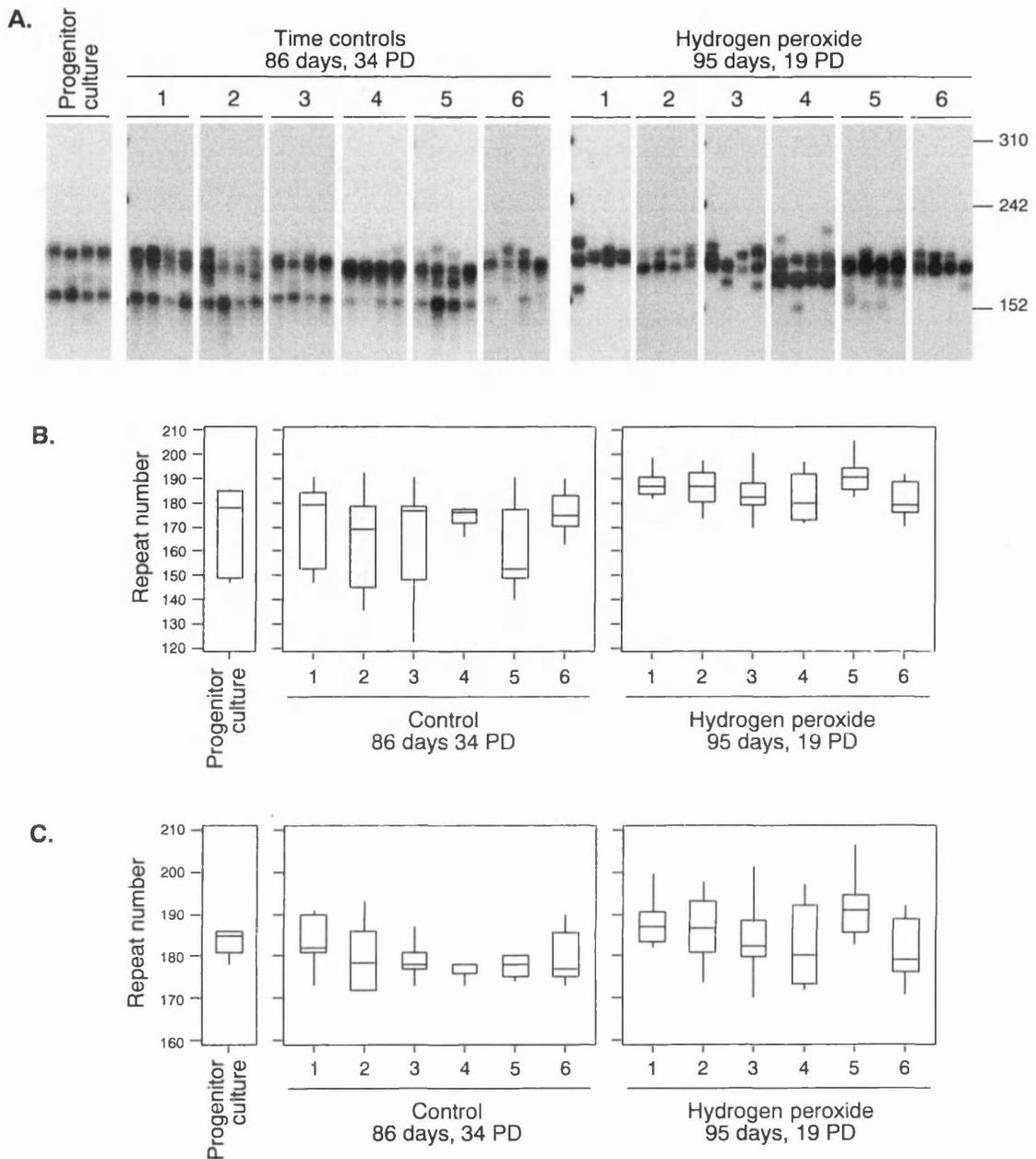
**Figure 5.4. Hydrogen peroxide treatment and expanded CAG-CTG repeat dynamics in D2763Kc2 cells.**

(A) The autoradiographs show representative SP-PCR amplifications of DNA samples extracted from replicate D2763Kc2 cells cultured for 80 population doublings with 100  $\mu\text{M}$  hydrogen peroxide (99 days). Two control cultures were also analysed: doublings control cells consisted of cultures maintained for 82 population doublings (97 days), and time control cultures were grown for 105 population doublings (122 days). The progenitor culture from which all cells were derived at day zero is shown on the left. The D2763Kc2 cell line was originally cloned by limiting dilution from D2763K kidney cells (see Figure 3.8). Fresh media, with or without hydrogen peroxide, was added to the cultures every 2-3 days and the cells were passaged weekly. An average of five to 30 transgene molecules were amplified in independent reactions. The molecular weight markers, converted into CTG repeat numbers, are displayed on the right. (B) The boxplots show the degree of repeat size variation observed in treated and control cultures. The top and bottom of the boxes correspond to the third (Q3) and first quartiles (Q1), respectively, and the line across the box displays the median repeat number. The lines extending from the top and the bottom of the boxes, include values that fall inside the lower and upper limits:  $Q1-1.5(Q3-Q1)$  and  $Q3+1.5(Q3-Q1)$ , respectively. The median rates of expansion, corrected for time and population doublings, were determined and revealed to be significantly different between control and treated cells ( $p < 0.05$ , two-tailed Mann-Whitney  $U$  test).

( $p=0.0202$ , two-tailed Mann-Whitney  $U$  test). Significant differences were also found when the median expansion rates were corrected for population doublings, with treated cells expanding 0.284 repeats per doubling, and the controls increasing their repeat number by 1.050 repeats per doubling ( $p=0.0131$ , two-tailed Mann-Whitney  $U$  test). However, it should be noted that hydrogen peroxide treated cells showed evidence of reduced size variability around sharp and well defined peaks of variability (Figure 5.4.B, black arrowheads), suggesting that some of this effect might have been mediated by clonal growth of cells selected for enhanced viability under conditions of oxidative stress, which were most certainly already present in the progenitor culture, although in such a small number that their detection by SP-PCR procedures was not possible.

### 5.2.3.2. D4132K kidney cell line

Similarly to what was previously described for novobiocin, the effect of hydrogen peroxide treatment on the dynamics of expanded CAG•CTG repeats was investigated in an additional kidney cell line that had only undergone a few population doublings (12 population doublings in 84 days) prior to hydrogen peroxide exposure, hence reducing the chances of accumulating multiple genetic mutations, and thereby resembling a primary cell culture. The analysis was performed on D4132K cells as described for the D2763Kc2 cell line, with the exception that the drug treatment was finished after 19 population doublings in culture (95 days), and the repeat size variability in treated cells was only compared with time controls. D4132K cultures experienced a 34% decrease in their proliferation capacity in the presence of the oxidant agent, as estimated by their population doubling time. The amplification of an average number of five to 20 molecules in multiple independent reactions revealed marked differences in the repeat size variability between treated and control cultures (Figure 5.5.A). The controls retained the bimodal repeat distribution previously detected in the progenitor culture, consisting of two major peaks of variability with median allele sizes of ~150 and ~180 repeats. In contrast, in treated cells the proportion of cells carrying shorter repeats decreased, and most of the replicate cultures were actually overtaken by the cells carrying longer alleles, with a median size of ~185 repeats. These results suggest that the addition of hydrogen peroxide to the growth medium might have greatly enhanced the repeat expansion rates of both subpopulations and/or induced a selection for cells displaying higher resistance to oxidative stress, which in this particular cell line happen to carry longer repeats. Since we cannot reject any of these two hypotheses based on our data, single molecule analysis was performed and two sets of comparisons were established (Figure 5.5.B). Regardless of the possibility of cell selection, expansion rates in hydrogen peroxide treated cells (0.0674 repeats per day) were compared with the overall repeat gain in control cells, considering both subpopulations altogether (-0.0279 repeats per day). The enhanced expansion-biased repeat instability detected in treated cells was highly significant ( $p=0.0082$ , two-tailed Mann-Whitney  $U$  test). In addition, in view of the great likelihood of cell selection in the presence of hydrogen peroxide, only the upper alleles detected in the progenitor culture and in the six time controls were included in the analysis, since they were most certainly favoured by the selection event that might have occurred in these cultures. Whereas



**Figure 5.5. Hydrogen peroxide treatment and expanded CAG·CTG repeat dynamics in D4132K kidney cells.**

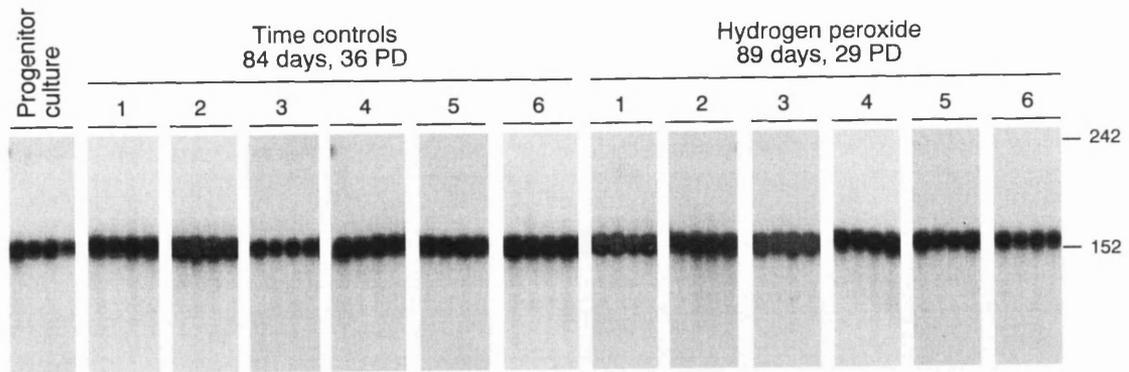
(A) The autoradiographs show representative SP-PCR analyses of DNA samples extracted from replicate D4132K cells cultured for 19 population doublings (PD) with 100  $\mu$ M hydrogen peroxide (95 days) as indicated in the legend to Figure 5.4. Time control cultures were maintained for 34 population doublings (86 days), and the progenitor culture from which all cells were derived at day zero were also analysed. An average of five to 20 transgene molecules were amplified in independent reactions. The molecular weight markers, converted into CTG repeat numbers, are displayed on the right. The boxplots show the degree of repeat length variability observed in treated and control cultures as described in Figure 5.4. Boxplots for the overall repeat size variability within control cultures (B), and for the upper alleles from which hydrogen treated cells might have derived (C) are presented individually. The median rates of expansion, corrected for time and population doublings, were determined following single molecule analysis, and revealed to be significantly different between control and treated cells, in both B and C analyses ( $p < 0.05$ , two-tailed Mann-Whitney *U* test).

control cells showed a median rate of contraction of 0.0808 repeats per day, hydrogen peroxide treated cells only exhibited a significantly lower median contraction of 0.00302 repeats per day ( $p=0.0131$ , two-tailed Mann-Whitney  $U$  test).

In conclusion, the addition of hydrogen peroxide to the cell growth medium appeared to alter the repeat dynamics in D4132K cell cultures, causing an apparent reduction in the rate of contraction of CAG•CTG repeats. The results were highly significant, even when the possible contribution of selection for a particular subpopulation of cells with enhanced viability was minimised by excluding the presumed unselected cells from the analysis. Nonetheless, a more subtle selection mechanism, within the subpopulation of cells carrying longer alleles, might still have accounted for the effect of hydrogen peroxide exposure on the dynamics of the transgenic CAG•CTG sequence in D4132K cells.

#### **5.2.4. Effect of hydrogen peroxide treatment on stable expanded CAG•CTG repeats**

The preliminary data presented in Section 5.2.2 suggested the reduced ability of oxidative stress to enhance the mutation rate repeat of stable CAG•CTG sequences. To further test this hypothesis the effect of the addition of hydrogen peroxide on the dynamics of long trinucleotide repeats was assessed in a kidney cell line derived from a *Dmt-E* mouse, previously reported to carry stable transgenic repetitive sequences (Section 3.2.9). The study was carried out as described before. The progenitor E3994K kidney culture had been growing for 81 days, and undergone 9 population doublings prior to the establishment of replicates and the beginning of the treatment. No repeat size variability was detected by SP-PCR analysis in the progenitor cells, which carried ~160 CAG•CTG repeats. Following a treatment period of 89 days (29 population doublings) with 100  $\mu$ M hydrogen peroxide, which increased the population doubling time by 31%, the repeat size variation was monitored by SP-PCR and compared to time control cultures, grown for 84 days (36 population doublings). More than 1000 transgene molecules were amplified in multiple independent reactions, with an average DNA input of ten to 30 transgene molecules per reaction. High DNA input SP-PCR amplifications failed to detect evident expansions or deletions within both control and hydrogen peroxide treated cell populations (Figure 5.6). The average repeat sizes were found to be identical in both sets of replicates and in the progenitor culture, indicating that the repeat remained remarkably stable, even under significant levels of oxidative stress. These results confirm that hydrogen peroxide does not induce detectable levels of repeat instability in cell lines carrying intrinsically stable trinucleotide sequences, and that *cis*-acting factors have greater impact on the control of repeat dynamics, compared to environmental agents, such as oxidative stress.



**Figure 5.6. Hydrogen peroxide treatment and CAG-CTG repeat dynamics in E3994K kidney cells.**

Six replicate cultures derived from an eight-month-old *Dmt-E* mouse (E3994K cell line) were treated with 100  $\mu$ M hydrogen peroxide for 89 days, corresponding to 29 population doublings (PD). Cells were treated and passaged as described in the legend to figure 5.4. The autoradiographs show representative SP-PCR analyses of DNA samples extracted at the end of the treatment. Time control cultures maintained for 84 days (36 population doublings), and the progenitor culture from which all the cells were derived at day zero are also shown. An average of ten to 30 transgene molecules were amplified in independent reactions. The molecular weight markers, converted into CTG repeat numbers, are displayed on the right. Similarly to the controls, the CAG-CTG transgenic repeats remained remarkably stable in treated cells, with no obvious expansions or deletions being detected by high DNA input SP-PCR analysis.

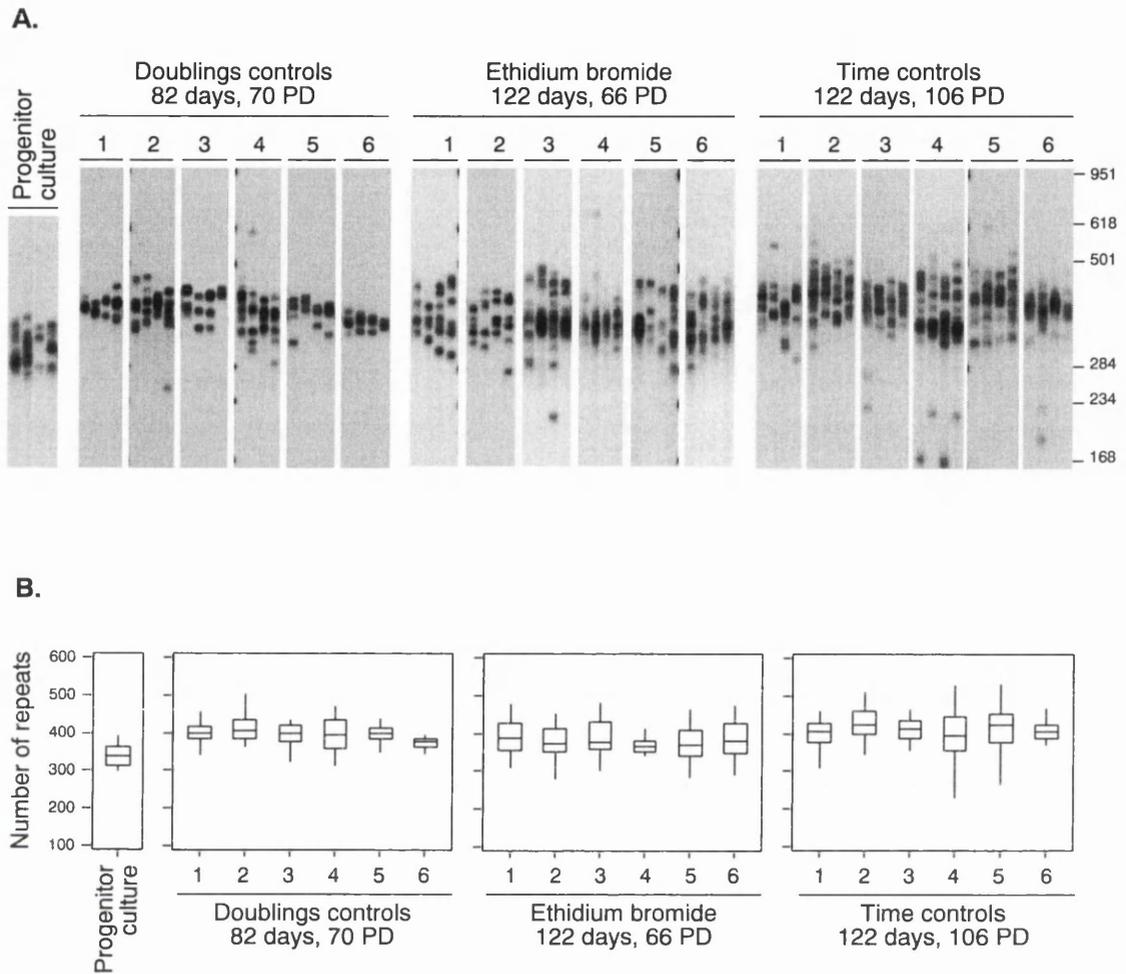
### 5.2.5. Triplet repeat instability in mouse cells with impaired mitochondrial function

The mitochondrial genome encodes for only 13 of more than one hundred mitochondrial proteins, the rest are encoded by the nucleus. Mitochondrial DNA depletion leads to increase ROS generation and cellular oxidative stress, probably because of incomplete biogenesis of the mitochondrial electron transport chain (Miranda *et al.*, 1999). A possible sequence of events resulting in severe mitochondrial dysfunction and oxidative stress could be: first, a functional defect in the mitochondrial electron transport chain due to improper protein assembly, or assembly of a protein complex with abnormal properties; second, diversion of electrons to form increased amounts of superoxide; third, induction of superoxide dismutase; fourth, increased formation of hydrogen peroxide from superoxide, via superoxide dismutase catalysis; fifth, generation of hydroxyl radicals from hydrogen peroxide, via the Fenton reaction or other mechanisms (Raha and Robinson, 2000). This severe scenario was recreated in *Dmt-D* cell cultures, and the outcome on the trinucleotide repeat metabolism investigated.

### 5.2.5.1. Effects of ethidium bromide on trinucleotide repeat dynamics

Ethidium bromide is a DNA intercalating drug, with known mutagenic properties (McCann *et al.*, 1975). Although ethidium bromide has been shown to be effective for isolating mitochondrial DNA-less ( $\rho^0$ ) cell lines from avian (Desjardins *et al.*, 1986) and human cells (Hayashi *et al.*, 1991; King and Attardi, 1989), it does not induce depletion of mitochondrial DNA in mouse cells. However, mitochondrial transcription is dramatically inhibited by ethidium bromide treatment (Inoue *et al.*, 1997a; Inoue *et al.*, 1997b; Morel *et al.*, 1999). Ethidium bromide is also capable of inhibiting endogenous mitochondrial respiration by reducing the activity of the respiratory complexes and ATPase activity (Miko and Chance, 1975), to generate point mutations in the mitochondrial genome (Morel *et al.*, 1999) and also to specifically inhibit mitochondrial protein synthesis by preventing polysome formation (Avadhani and Rutman, 1975). For all these reasons exposure to ethidium bromide is thought to result in dramatic mitochondrial dysfunction and increased levels of intracellular oxidative stress.

To investigate the effect of ethidium bromide-induced oxidative stress in the dynamics of expanded CAG•CTG trinucleotide repeats, six replicates established from the D2763Kc2 clonal cell line were treated with 250 nM of ethidium bromide for 122 days. During this period of time the cultures underwent 66 population doublings, corresponding to a reduction of 66% in their proliferative capacity, compared to the controls. SP-PCR procedures allowed the analysis of repeat size variability in both treated and control cells (Figure 5.7.A). The analysis indicated continuing trinucleotide expansion in all sets of cultures, but revealed a decreased expansion rate in cells exposed to ethidium bromide, compared with the controls, particularly the time controls. More interestingly, the reduction in repeat expansion shows no association with clonal proliferation of selected cells as reported for hydrogen peroxide treatments. To quantify the effect of the drug treatments on the stability of trinucleotide repeats in cultured cells, single molecule analysis was performed (Figure 5.7.B) and the expansion rate of the CAG•CTG repeats compared between treated and control cells, as described previously. The control cells showed a median repeat gain of 0.615 repeats per day, in contrast with the median expansion of 0.310 repeats per day exhibited by ethidium bromide treated cells ( $p=0.0051$ , two-tailed Mann-Whitney  $U$  test). Similarly, statistically significant differences were also revealed when the expansion rates were corrected for population turnover: median gain of 0.890 and 0.602 repeats per population doubling in control cells and treated cells, respectively ( $p=0.0202$ , two-tailed Mann-Whitney  $U$  test). In summary, ethidium bromide treatment resulted in decreased rates of expansion, with no association with the accumulation of cells carrying large deletions or reduction in repeat size variability, characteristic of a selection process similar to that observed following exposure to hydrogen peroxide.

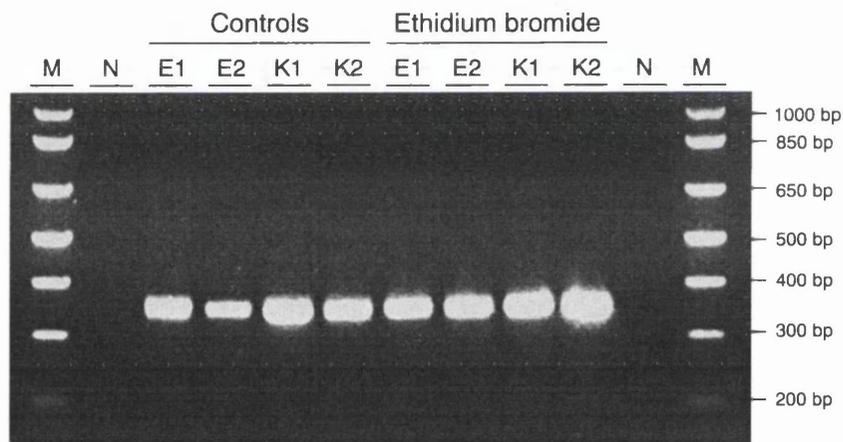


**Figure 5.7. Ethidium bromide treatment and expanded CAG·CTG repeat dynamics in D2763Kc2 cells.**

(A) The autoradiographs show representative SP-PCR amplifications of DNA samples extracted from replicate D2763Kc2 cells cultured for 66 population doublings (122 days) with 250 nM ethidium bromide. Doublings control cells were maintained for 70 population doublings (82 days), whereas time controls were grown for 122 days (106 doublings). The progenitor culture from which all cells were derived at day zero is shown on the left. Fresh media, with or without ethidium bromide, was added to the cultures every 2-3 days and the cells were passaged weekly. An average of five to 20 transgene molecules were amplified in independent reactions. The molecular weight markers were converted into CTG repeat numbers, and shown on the right. (B) The boxplots show the degree of repeat size variability observed within treated and control cultures. The median rates of expansion, corrected for time and population doublings, were determined for treated cells, and compared with both controls. The results proved to be significantly different between control and treated cells ( $p < 0.05$ , two-tailed Mann-Whitney  $U$  test).

### 5.2.5.2. Investigating the biological bases of the effects of ethidium bromide on trinucleotide repeat dynamics

To gain insight into the mechanism(s) that mediate(s) the effect of ethidium bromide on trinucleotide repeat dynamics, the presence of mitochondrial DNA in cells chronically exposed to the chemical was assessed by PCR amplification of a 355-bp mitochondrial DNA sequence, which maps within the *12S ribosomal RNA* gene. The product was amplified with oligonucleotide primers H1 and L1 (Table 2.7) from DNA samples collected from D2763K cells grown either with or without ethidium bromide added to the medium (Figure 5.8). All DNA samples generated great amounts of a PCR product with the predicted size. The results could not be considered for quantitative densitometric analysis, given the large excess of mitochondrial DNA used as template in PCR amplification. Therefore, we can only suggest that continuous exposure of *Dmt-D* mouse cells to ethidium bromide did not result in gross depletion of the mitochondrial genome. Subtle differences, if any, could only be detected by reducing the amount of input mitochondrial DNA, by using competitive PCR techniques or by including a nuclear locus in the analysis, as an internal reference for mitochondrial DNA quantification.



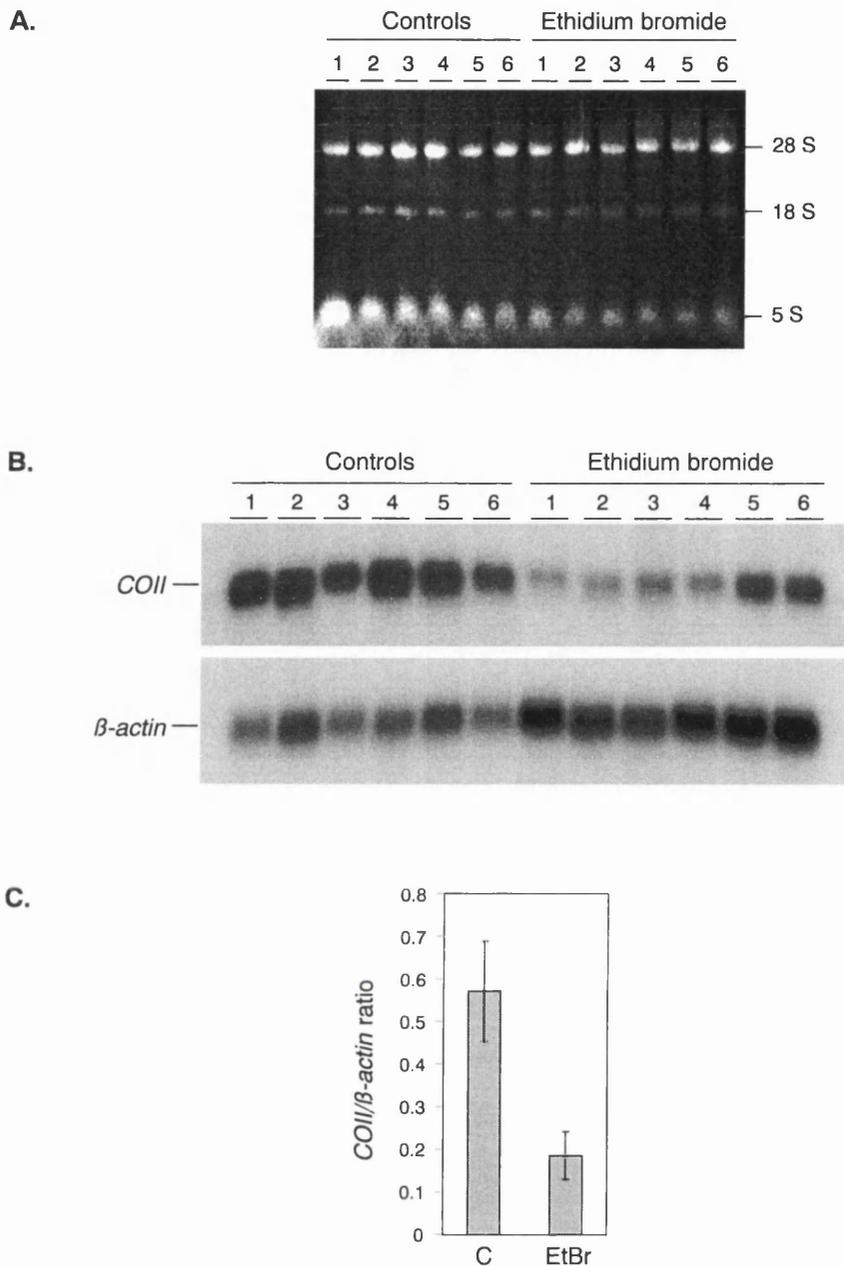
**Figure 5.8. PCR amplification of mouse mitochondrial DNA extracted from *Dmt-D* cells treated with ethidium bromide.**

In order to determine the effect of ethidium bromide exposure on the integrity of mouse mitochondrial genome, DNA samples collected from mouse cultured cells were amplified with oligonucleotide primers L1 and H1, which amplify a 355-bp sequence within the mitochondrial *12S rRNA* gene. Two control replicate cultures were maintained under standard growth conditions, in the absence of ethidium bromide, whereas treated replicates were exposed to 250 nM (kidney cells, K1 and K2) or to 1  $\mu$ M (eye cells, E1 and E2) of the chemical for at least 50 generations. Ten to 100 ng of total DNA were amplified by PCR, using 62°C as annealing temperature, and 30 cycles of amplification (Section 2.5.5). The PCR products were electrophoresed through an ethidium bromide-stained 1.5% (w/v) agarose gel (Section 2.5.7.1). The scale on the right represents the molecular size markers (M). Lanes N represent no template DNA controls.

To confirm the effect of ethidium bromide on mitochondrial DNA transcription, the steady-state transcript levels of the mitochondrial encoded enzyme cytochrome oxidase II were investigated in treated cells. Total RNA was extracted from six D2763Kc2 cultures treated with 250 nM ethidium bromide for 95 days, and six control cultures. The integrity of the RNA samples was verified by agarose gel electrophoresis, and the analysis revealed the presence of good quality RNA (Figure 5.9.A). Northern blot analysis was performed in order to quantify the levels of mitochondrial *cytochrome oxidase II* transcripts against nuclear-encoded  $\beta$ -actin mRNA levels (Figure 5.9.B). Densitometric analysis was carried out, taking care to perform the quantification in the linear range of signal intensity (Figure 5.9.C). A highly significant reduction by  $\sim 2/3$  in the mRNA levels of *cytochrome oxidase II* in exposed cells compared with the controls was found ( $p=0.0002$ , two-tailed  $t$ -test), confirming that ethidium bromide strongly affects mitochondrial transcription, and most certainly causes great levels of oxidative stress, which may therefore mediate the changes in repeat dynamics.

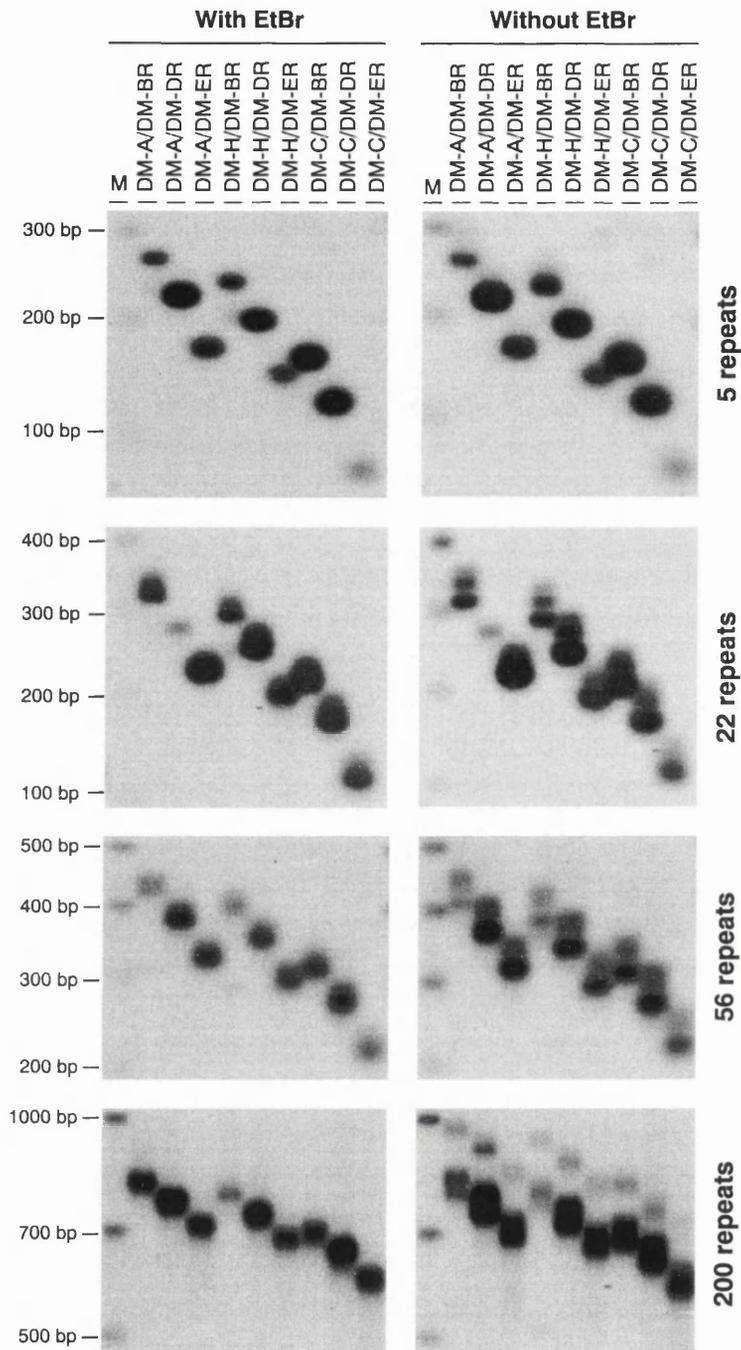
### 5.2.5.3. Altered electrophoretic mobility of trinucleotide repeat PCR products in the presence of ethidium bromide

In addition to investigating the biological consequences of ethidium bromide exposure on mitochondrial metabolism, the intercalating properties of this chemical and the possible effect upon the complex conformational fluctuations and physical features of trinucleotide repeats were also analysed. PCR products containing different sized CAG•CTG tracts were amplified from the human *DM1* locus (as described in Section 4.2.1.1) with various oligonucleotide primer combinations. The samples were electrophoresed through agarose gels, either with or without 500 nM ethidium bromide, in both gel and electrophoresis buffer, and detected by Southern “squash” blot hybridisation (Figure 5.10). In the presence of ethidium bromide, all PCR products migrated as a single DNA species, and generated a sharp and well defined band. In contrast, amplified repetitive sequences containing 22, 56 and 200 triplet repeats generated multiple alternative bands in the absence of ethidium bromide. The major bands observed in the absence of the intercalating chemical exhibited the same electrophoretic mobility as the single band detected when ethidium bromide was added to the gel and buffer. However, the additional DNA species exhibited apparent sizes that were  $\sim 10\%$  higher for PCR products containing 22 and 56 repeats, and  $\sim 20\%$  higher for the amplification products carrying 200 CAG•CTG units. In addition to the enhanced gel retardation for the additional DNA species, the longest PCR products also generated a couple of fast migrating bands, instead of a single one (Figure 5.10). In other words, not only the intensity of the effect of ethidium bromide on the mobility of the PCR products, but also the complexity of the electrophoretic profile in the absence of the intercalating chemical, appear to increase as the repetitive tract gets longer, suggesting that the addition of ethidium bromide to the gel and buffer must change the mobility of structural isomers formed within CAG•CTG expanded sequences. Interestingly, PCR products containing only five CAG•CTG repeats did not generate detectable additional DNA bands in the absence of ethidium bromide, although the same products have been



**Figure 5.9. Quantification of cytochrome oxidase II mRNA levels in ethidium bromide treated cells.**

(A) RNA visualisation by agarose gel electrophoresis to check the integrity and the quality of the RNA samples extracted from six D2763Kc2 replicate cultures treated with 250 nM ethidium bromide for 95 days, and six control cultures. Note the labelled bands corresponding to rRNA species 28S, 18S and 5S. (B) The autoradiographs show a northern blot analysis of mRNA levels of *cytochrome oxidase II* (*COII*) and *β-actin* transcripts in ethidium bromide treated cells, compared to control cultures. See Sections 2.2.6 and 2.6.4 for further details. (C) The graph shows the quantitative analysis of expression levels of *cytochrome oxidase II* relative to *β-actin* nuclear control, in both control (C) and ethidium-bromide treated (EtBr) cultures. The observed reduction in *cytochrome oxidase II* mRNA levels in treated cells is highly significant ( $p=0.0002$ , two tailed *t*-test).

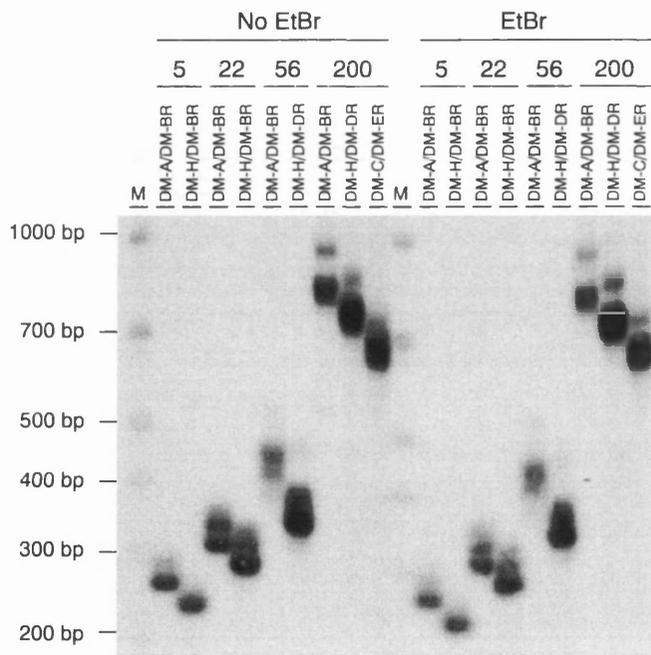


**Figure 5.10. Effect of ethidium bromide on the mobility of expanded CAG•CTG sequences in agarose gels.**

PCR products containing five, 22, 56 or 200 CAG•CTG repeats were amplified from the human *DM1* locus with various combinations of oligonucleotide primers (DM-A, DM-H, DM-C, DM-BR, DM-DR and DM-ER) and resolved through 1.8% (w/v) agarose gels, with (left) or without (right) 500 nM ethidium bromide in both the gel and running buffer. The amplified products were subsequently detected by Southern “squash” blot hybridisation. The scale on the left shows the position of the molecular weight markers (M) in base pairs. The autoradiographs illustrate the increased mobility of alternative expanded CAG•CTG conformers in agarose gels in the presence of ethidium bromide. Note the multiple alternative products observed in the absence of the intercalating chemical.

shown to have the ability to give rise to, at least, some alternative structures *in vitro*, as demonstrated by native PAGE (Section 4.2.1.1).

Surprisingly, when the PCR products were incubated with 500 nM ethidium bromide for 48 hours prior to electrophoresis, they still generated additional slow-migrating bands when resolved through an agarose gel without ethidium bromide (Figure 5.11). This observation suggests that ethidium bromide does not interconvert stable S-DNA structures. Given that the presence of ethidium bromide is required in the gel and electrophoresis buffer so that a single band is detected by Southern “squash” blot hybridisation, it may be hypothesised that this chemical interferes with the migration of alternative DNA structures adopted by CAG•CTG repetitive sequences through agarose gels. Imagining that ethidium bromide is able to bind to DNA over a 48-hour incubation period, it is possible that it is removed from the DNA samples during the migration through the gel, since DNA and ethidium bromide are forced to move in opposite directions through electrophoresis. This might explain the generation of slow migrating bands by PCR products treated with ethidium bromide.



**Figure 5.11. Effect of ethidium bromide incubation on the mobility of expanded CAG•CTG sequences through agarose gels.**

Human DNA sequences derived from the *DM1* locus, and carrying different repeat numbers were amplified by PCR using multiple oligonucleotide primer combinations, indicated in the figure above each lane. Half of each PCR product was incubated with 500 nM of ethidium bromide for over 48 hours prior to electrophoresis through a 1.8% (w/v) agarose gel and detected by Southern “squash” blot hybridisation. The autoradiographs do not reveal any obvious difference in the electrophoretic profile between treated and non-treated samples. The scale on the left represents the position and sizes (in base pairs) of the molecular weight markers (M).

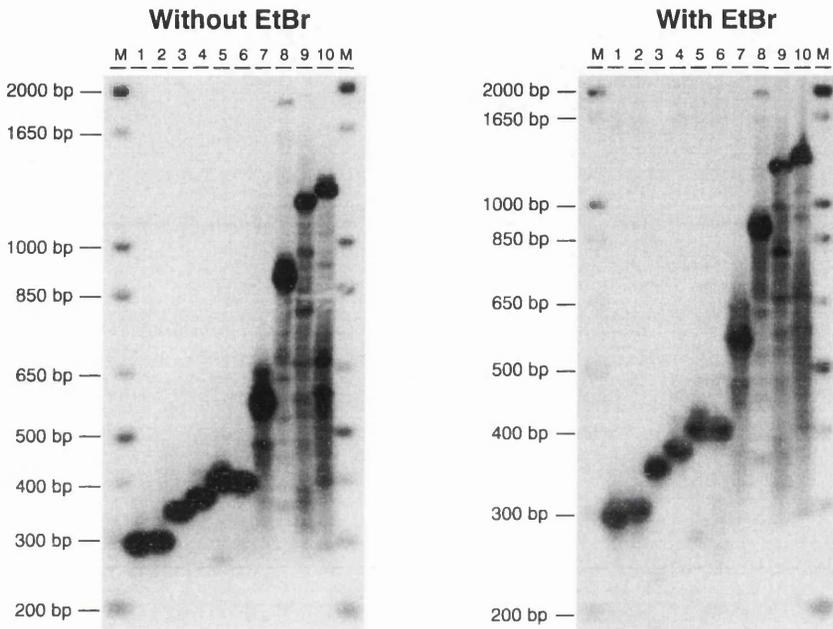
#### 5.2.5.4. Searching for alternative DNA structures *in vivo*

Despite the simplicity of the assay, the findings described in the last section opened new avenues for the study of alternative DNA structures adopted by trinucleotide repeat sequences *in vivo*. Given that ethidium bromide affects the electrophoretic mobility of conformational isomers formed within repetitive CAG•CTG sequences, genomic DNA samples were analysed to test if alternative structures might exist *in vivo*. The hypothesis being that if different electrophoretic profiles were observed in the absence of ethidium bromide, this would be a preliminary indication that alternative DNA conformations may also occur *in vivo*, and also that the effect of ethidium bromide on the repeat dynamics in cell culture might be a consequence of a modified DNA topology induced by exposure to the intercalating dye.

The presence of alternative slow-migrating DNA conformations was initially assessed in bacteria. Different sized CAG•CTG sequences, ranging from 22 to 450 repeats, were cloned into pGEM<sup>®</sup>-T Easy bacterial vector, and transformed into *E. coli* (for further details see Section 2.2.4). The plasmid DNA was purified from a 5-ml culture, grown overnight, and digested with *EcoRI* restriction endonuclease, prior to electrophoretic analysis, either with or without 500 nM ethidium bromide in the gel and electrophoresis buffer, and finally detected by Southern “squash” blot hybridisation (Figure 5.12). The electrophoretic mobility profiles in both autoradiographs are identical, not only in terms of the number of bands detected but also in the mobility of each band relative to the molecular weight markers, suggesting that the presence or absence of ethidium bromide in the gel and buffer does not reveal a difference in the mobility of CAG•CTG-containing plasmid DNA, purified from *E. coli*. It is worth mentioning the dramatic deletion-biased instability, particularly for the larger trinucleotide repeat tracts (Figure 5.12), already described for microbial model systems (Freudenreich *et al.*, 1997; Kang *et al.*, 1995b; Schweitzer and Livingston, 1997; Wells *et al.*, 1998).

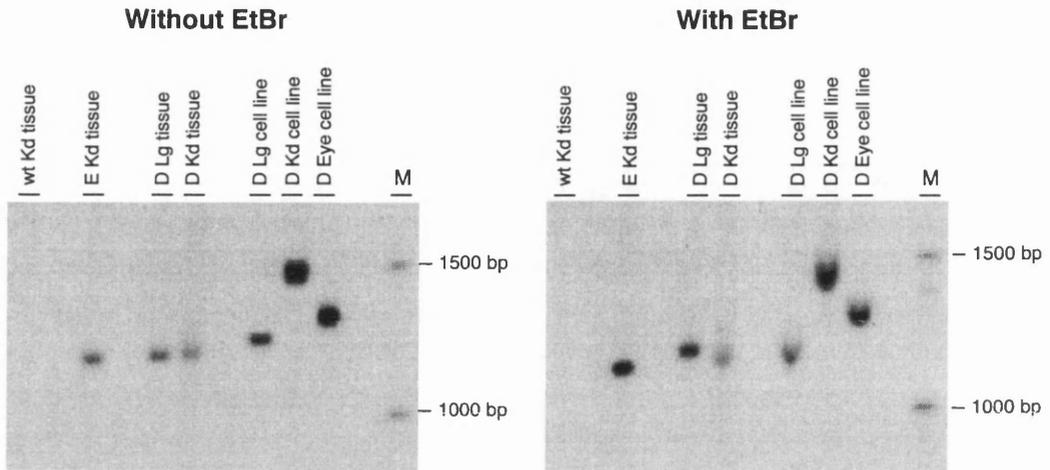
Interestingly, although bacterial plasmid DNA, containing CAG•CTG repeats, is capable of generating alternative S-DNA structures, when subjected to a reannealing protocol, no evidence of alternative DNA conformations was reported for purified plasmid DNA, following native PAGE (Pearson *et al.*, 1998b).

To address the possibility that alternative structures could be detected by this method in genomic DNA samples purified from *Dmt-D* mouse tissue or cultured cells, Southern blot hybridisation techniques were performed following agarose gel electrophoresis in the presence or absence of ethidium bromide added to the gel and buffer. As observed with purified plasmid DNA, no additional low mobility bands were observed in the absence of ethidium bromide in purified mouse genomic DNA samples (Figure 5.13). Some minor differences in the mobility of a few bands were however detected between the two autoradiographs shown, but they most likely resulted from differences in salt concentration in the sample, and are therefore not relevant for this study (Figure 5.13). Nevertheless, a minor increase in the electrophoretic mobility of mouse genomic DNA fragments (up to 5%) was detected in the presence of ethidium bromide, relative to the molecular weight markers.



**Figure 5.12. Effect of ethidium bromide on the mobility of CAG•CTG-containing plasmid DNA through agarose gels.**

CAG•CTG repeat sequences varying from 22 to 160 units in size were amplified with DM-A and DM-DR oligonucleotide primers from the human *DM1* locus. The PCR products were ligated into pGEM®-T Easy and transformed into *E. coli* TOP10 bacterial strain. Plasmids purified from 5 ml cultures grown overnight were digested with *EcoRI* and resolved through 1.8% (w/v) agarose gels with (right) or without (left) 500 nM ethidium bromide (EtBr) in both the gel and running buffer, and detected by Southern “squash” hybridisation (lanes 1-7). pGEM-T750.19, pGEM-T750.21 and pGEM-T750.22 constructs, carrying longer CAG•CTG sequences (between 250 and 450 repeats) amplified from the human *DM1* locus with oligonucleotide primers DM-H and DM-BR (Section 2.2.4), were analysed as described above (lanes 8-10). The scale on the left displays the sizes of the molecular weight markers (M) in base pairs. No major differences in the mobility profiles were detected between the two gels, independently of the presence or absence of ethidium bromide. Note the preferential accumulation of deletion mutants, particularly for longer trinucleotide repeat tracts (lanes 7-10).



**Figure 5.13. Effect of ethidium bromide on the mobility of CAG•CTG-containing mouse genomic DNA fragments.**

Mouse genomic DNA samples (20  $\mu$ g) were digested with 40 units of *Eco*RI for 16 hours and the digested products resolved through 1% (w/v) agarose gels, with (right) or without (left) 500 nM ethidium bromide (EtBr) in both the gel and running buffer. The CAG•CTG-containing transgenic sequences were detected by Southern hybridisation with radiolabelled DM-F/DM-PRENK probe (Section 2.2.6), which maps downstream to the repetitive tract. Genomic DNA samples were extracted from the following sources: kidney of a wild-type mouse (lane “wt Kd tissue”), *Dmt*-E kidney tissue sample (lane “E Kd tissue”), *Dmt*-D lung tissue sample (lane “D Lg tissue”), *Dmt*-D kidney tissue sample (lane “D Kd tissue”), *Dmt*-D lung cell line (lane “D Lg cell line”), *Dmt*-D kidney cell line (lane “D Kd cell line”) and *Dmt*-D eye cell line (lane “D Eye cell line”). The scale on the right represents the molecular weights of the size markers (M) in base pairs. Multiple alternative conformers were not observed in the absence of ethidium bromide. Yet, a slight increased mobility (up to 5%) was detected in the presence of ethidium bromide, relative to the molecular weight markers.

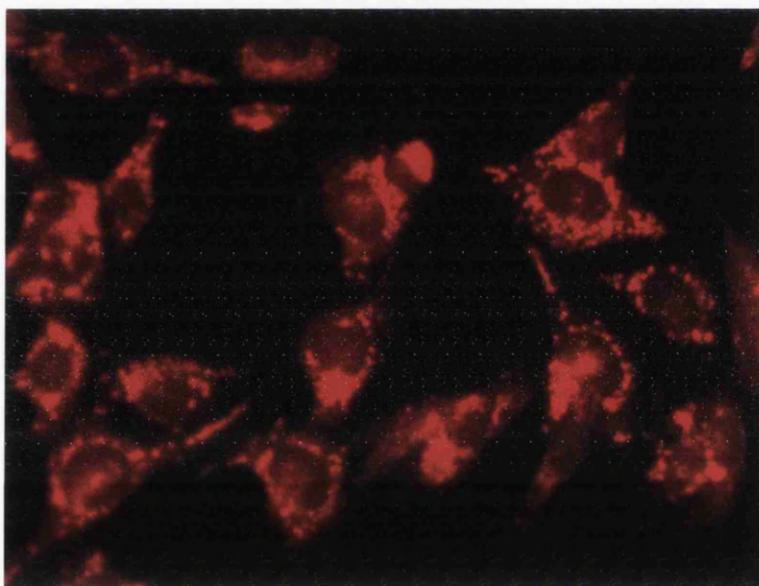
If alternative DNA structures do occur *in vivo*, both in bacteria and in mammalian cells, they may not exist at sufficient levels to be detected by standard Southern blot hybridisation in the absence of ethidium bromide. Alternatively, DNA purification protocols may introduce nicks and single strand breaks, which may cause relaxation of non-B-DNA conformations.

#### 5.2.5.5. Effects of rhodamine-6G on trinucleotide repeat dynamics

Inhibition of mitochondrial electron transport chain may accelerate oxidant production (Duranteau *et al.*, 1998). The lipophilic dye rhodamine-6G binds very tightly to mitochondria, and causes powerful inhibition of oxidative phosphorylation, drastically reducing the efficiency of ADP phosphorylation without damaging the mitochondrial structure (Gear, 1974). Great levels of oxidative stress are therefore expected following exposure to rhodamine-6G. Moreover, rhodamine-6G has not been reported to interact directly with nuclear genomic DNA, and any effect derived from exposure to this chemical might be considered to result from an overproduction of intracellular ROS by the mitochondria. This drug was therefore selected to confirm the hypothesis that oxidative stress may be a mediator of repeat metabolism, affecting the mutation rate of trinucleotide repetitive sequences.

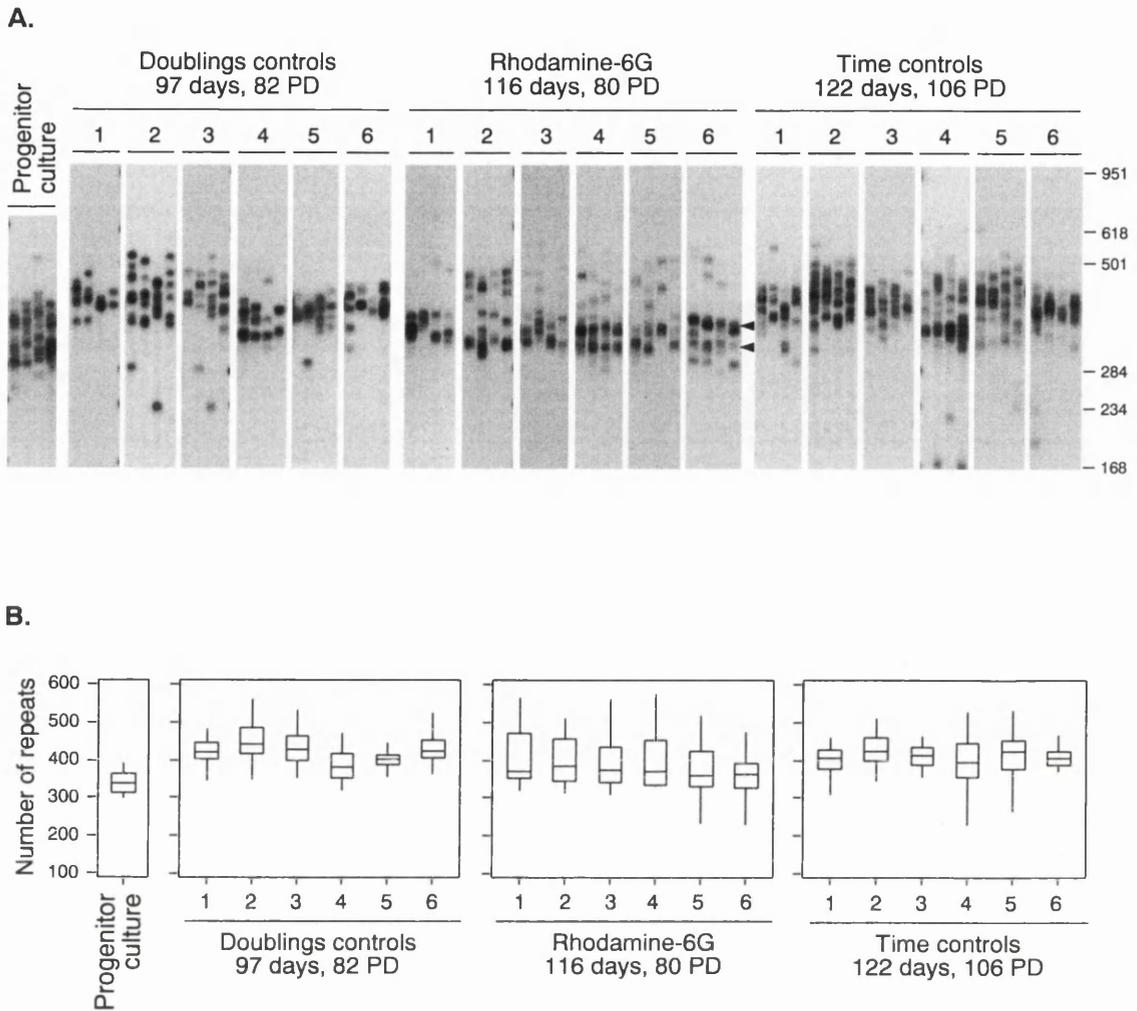
D2763Kc2 kidney cells were treated with 50 nM rhodamine-6G for ten days and the intracellular localisation of drug analysed by fluorescence microscopy (Figure 5.14). Rhodamine-6G staining was detected in numerous cytoplasmic foci, particularly in the nuclear periphery, which is consistent with mitochondrial accumulation of the drug. Furthermore, an increase of 22% in the population doubling time was estimated, in agreement with enhanced levels of oxidative stress. Following a 116-day exposure period to 50 nM rhodamine-6G, corresponding to 80 population doublings, the trinucleotide repeat size variability was assessed in six D2763Kc2 replicate cultures by SP-PCR procedures (Figure 5.15.A). As described for hydrogen peroxide treated cells, the analysis revealed an overall decrease in the average repeat size in cells exposed to rhodamine-6G, with evidence for cell selection, mediated by growth advantage of mutant cells displaying enhanced resistant to the chemical (Figure 5.15.A, black arrowheads). Single molecule analysis was carried out, and the expansion rates, corrected for both time and population doublings, were calculated and compared between treated and control cells, as described previously (Figure 5.15.B). Whereas the control cells exhibited a median rate of expansion of 0.615 repeats per day, rhodamine-6G treated cells only gained a median number of 0.298 repeats per day ( $p=0.0051$ , two-tailed Mann-Whitney  $U$  test). A decrease in the median repeat expansion from 1.050 to 0.432 repeats per population doubling was also found to be statistically significant between control and treated cultures, respectively ( $p=0.0082$ , two-tailed Mann-Whitney  $U$  test).

Despite the lower median repeat number in cells treated with rhodamine-6G, and evidence for clonal expansion driven by selection, at least in some of the exposed replicates, a large proportion of cells carrying longer alleles, containing ~400-600 repeats, was detected in treated cultures. In fact, the variance for the allele sizes carried by cells exposed to rhodamine-6G was significantly higher than in the six doubling control cultures ( $p=0.0082$ , two-tailed Mann-Whitney



**Figure 5.14. Intracellular localisation of rhodamine-6G in mouse kidney cells.**

The picture reveals the intracellular localisation of cultured mouse cells treated with 50 nM rhodamine-6G for ten days. The chemical accumulates in multiple cytoplasmic foci, mainly around the nucleus, which are likely to represent mitochondria.



**Figure 5.15. Rhodamine-6G treatment and expanded CAG-CTG repeat dynamics in D2763Kc2 cells.**

(A) Representative SP-PCR amplifications of DNA samples extracted from replicate D2763Kc2 cells are shown in the autoradiographs. D2763Kc2 cells were cultured for 80 population doublings (116 days) with 50 nM rhodamine-6G. Doublings control cells were maintained for 82 population doublings (97 days), and time controls were grown for 122 days (106 doublings), both in the absence of rhodamine-6G. The progenitor culture, from which all cells were derived at day zero is shown on the left. Fresh media, with or without rhodamine-6G, was added to the cultures every 2-3 days and the cells were passaged weekly. An average of five to 30 transgene molecules were amplified in independent reactions. The molecular weight markers, converted into CTG repeat number, are shown on the right. (B) The boxplots show the degree of repeat size variability observed within treated and control cultures. The median rates of expansion, corrected for time and population doublings, were determined for treated cells, and compared with the closest control in terms of time and population doublings, respectively. The treatments resulted in significant differences in the rate of expansion between treated cells and both controls ( $p < 0.05$ , two-tailed Mann-Whitney  $U$  test).

*U* test), and five of the time control replicates ( $p=0.0358$ , two-tailed Mann-Whitney *U* test). Therefore, one may assume that an initial selection for cells, exhibiting higher resistance to rhodamine-6G, took place over the initial stage of the treatment, followed by an increased mutation rate as a result of the great levels of oxidative stress. To test this hypothesis, the repeat size variability at the end of the rhodamine-6G treatment was compared with the repeat size variation at an intermediate time point (Figure 5.16.A). None of the cultures revealed a significant decrease in the median repeat size over the second half of the treatment, rejecting a continuous selection for cells carrying shorter repeats (Figure 5.16.B). In fact, one of the cultures (replicate 1) actually exhibited a significant increase in the median repeat number, from 341 to 371 ( $p=0.0042$ , two-tailed Mann-Whitney *U* test). However, only replicates 1 and 2 showed a significant increase in the variation of the allele sizes between 40 to 80 population doublings under exposure to rhodamine-6G ( $p<0.05$ , *F*-test).

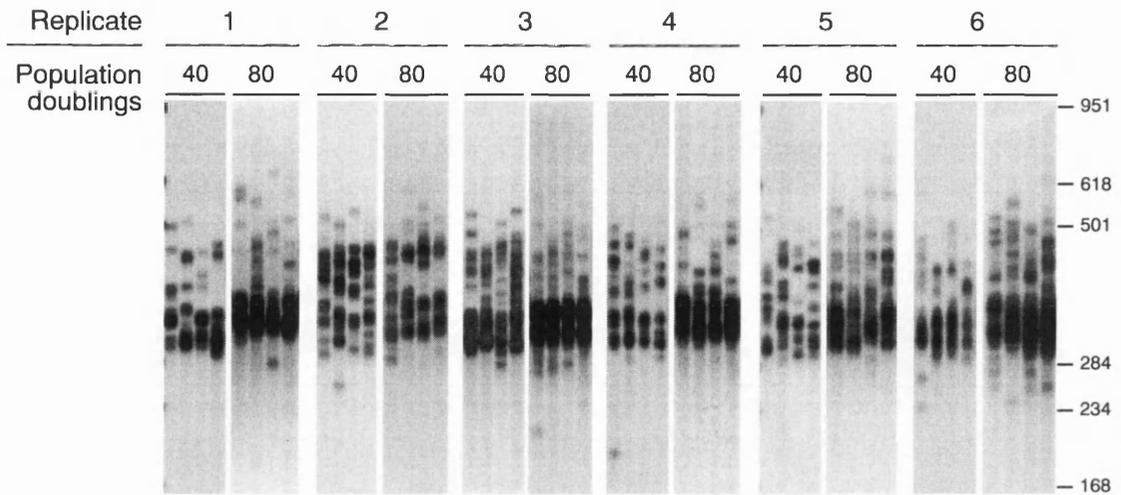
These results indicate that the possible selection process, which caused a significant reduction in the median allele length, was most certainly an early event that occurred shortly after the first exposure to rhodamine-6G, or at least over the first 40 population doublings. A continuing reduction in repeat size is not likely to have occurred. However, the data failed to indicate an increased accumulation of expanded mutant alleles following the initial expansion of viable cells.

### 5.3. Discussion

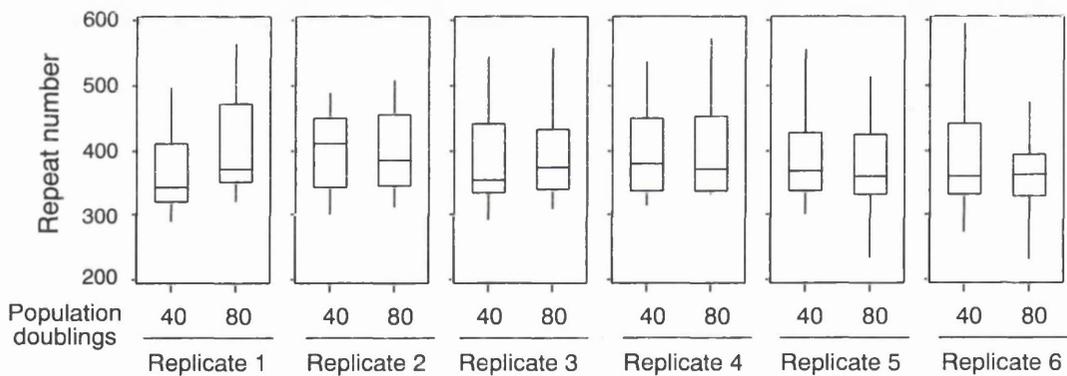
DNA damage caused by exposure to ROS, such as hydroxyl radicals, singlet oxygen, superoxide radical and hydrogen peroxide, is believed to be one of the most important reasons for endogenous DNA decay and spontaneous mutation (Ames *et al.*, 1993; Beckman and Ames, 1997; Rossman and Goncharova, 1998). During normal cell metabolism, the genome is exposed to numerous reactive oxidative substances. A range of endogenous processes has been assessed for their ability to generate DNA reactive species, such as lipid peroxidation products, endogenous alkylating agents, glycoxidation products, estrogens, reactive nitrogen species, chlorinating reagents, heme precursors and amino acids (Burcham, 1999). Most of these species might attack DNA bases or sugars and cause critical damage and may affect genomic stability.

To investigate a possible association between the cellular redox status and the dynamics of triplet repeats, *Dmt-D* cell lines were used during the course of this work. The susceptibility of two different cell lines to oxidative stress was assessed and correlated with the levels of somatic mosaicism previously detected. Higher cellular sensitivity to hydrogen peroxide in culture matched the greater levels of trinucleotide repeat instability exhibited by one of the two cell lines investigated. The findings presented here are indeed consistent with a parallel between cellular sensitivity to oxidative stress and greater levels of trinucleotide repeat dynamics, but an extended analysis of multiple cell lines, showing different levels of somatic mosaicism, should be performed to determine the statistical significance of this correlation.

**A.**



**B.**



**Figure 5.16. Investigating the progressive effect of rhodamine-6G on expanded CAG-CTG repeat dynamics.**

(A) The autoradiographs show representative SP-PCR amplifications of DNA samples collected from six replicate D2763Kc2 cells following treatment with 50 nM rhodamine-6G for 40 and 80 population doublings (PD). An average of ten to 100 transgene molecules were amplified in independent reactions. The molecular weight markers were converted into CTG repeat numbers, and displayed on the right. (B) The boxplots show the degree of repeat size variability and the median repeat number for each culture at both time points. None of the cultures showed a statistically significant decrease in the median repeat size between 40 and 80 population doublings. Replicate 1 exhibited a significant increase in the median repeat size over the same period of time ( $p=0.0042$ , two-tailed Mann-Whitney  $U$ -test), with an expansion rate of 0.762 repeats per population doubling, or 0.347 repeats per day.

Although the activity of enzymes directed to eliminate ROS varies considerably between cell types and changes with cell growth state, the sensitivity to oxidative injury does not reflect the levels of detoxifying enzymes (such as catalase and superoxide dismutase) in cultured cells (Duthie and Collins, 1997). The mutagenic response of mammalian cells to hydrogen peroxide is more likely to depend on the extent of free radical generation, and the inherent DNA repair activity rate and fidelity, and finally cell growth (Duthie and Collins, 1997; Henle and Linn, 1997; Rossman and Goncharova, 1998). For instance, lipopolysaccharide-induced oxidative stress increases the BER activity in extracts prepared from mouse monocytes and fibroblasts, and consequently their resistance to oxidative stress (Chen *et al.*, 1998a). Assuming that *Dmt-D* lung cells exhibit greater antioxidant defences, one may speculate that these cultures exhibit lower levels of intracellular ROS, and therefore lower spontaneous mutation frequency and a lower mutation rate within the transgenic trinucleotide sequence. In contrast, the higher resistance of lung cells to hydrogen peroxide might alternatively result from the lower activity levels of the DNA repair enzymes exhibited by this cell line, thereby avoiding the apoptotic pathway. Human cells selected for their resistance to hydrogen peroxide display reduced OGG1 and MYH activity and increased levels of anti-apoptotic protein BCL-2, consistent with the development of a resistance mechanism that bypasses apoptosis (Gu *et al.*, 2001). In line with this view, the higher DNA repair activity level exhibited by *Dmt-D* kidney cells, which eventually leads to apoptosis and to a greater death rate under oxidative stress compared to the lung cell line, may contribute to the increased levels of repeat instability in this cell line, according to recent models of trinucleotide mutation (Manley *et al.*, 1999b; van Den Broek *et al.*, 2002, Chapter 8).

In order to test for the ability of oxidative stress to modify trinucleotide repeat dynamics, a *Dmt-D* kidney clonal cell line was exposed to hydrogen peroxide, ethidium bromide, and rhodamine-6G. Whereas hydrogen peroxide is capable of originating extremely reactive hydroxyl radicals by the Fenton reaction (Henle and Linn, 1997), ethidium bromide and rhodamine-6G cause mitochondrial dysfunction and consequent production of ROS (Avadhani and Rutman, 1975; Gear, 1974; Miko and Chance, 1975; Morel *et al.*, 1999). In any case, increased levels of intracellular oxygen species are expected to cause great levels of oxidative stress. Whatever the treatment, a decreased rate of repeat expansion was observed between treated and control cells, suggesting that trinucleotide repeats show higher stability under particular conditions of slow cell proliferation induced by oxidative stress. Some of the hydrogen peroxide and rhodamine-6G treated cultures exhibited clear evidence of reduced repeat length variability, indicating that some of the effects observed might have been mediated by clonal expansion of cells selected for enhanced viability under conditions of oxidative stress, which happened to carry shorter repeats than the average repeat length in the overall population.

To corroborate the establishment of a selection process in culture under oxidative stress, the opposite effect on the expansion rate of trinucleotide repeats was observed following the exposure of D4132K kidney cells to hydrogen peroxide. The median repeat length of the transgenic sequence was significantly higher in D4132K treated cells, compared to the controls. This result appears to suggest that increased levels of oxidative stress may result in either an increase or a

decrease of the median repeat length of the CAG•CTG tracts, depending on the nature of the progenitor culture selected to carry the experiment. The bidirectionality of this phenomenon strongly suggests that the outcome of the exposure to great levels of oxidative stress on the trinucleotide dynamics in *Dmt-D* cultures cells is most certainly the result of a selection event, which favours cells exhibiting enhanced resistance to oxidative stress, rather than the consequence of a direct effect of ROS on the mutation rate of trinucleotide repetitive sequences.

It should be stressed at this point that a major component of oxidative stress defences is to keep actual cellular oxygen tension to a minimum. Oxygen concentration decreases from atmospheric levels (20%) in the lungs to only 2-5% in the tissues. Therefore, cultured cells grown in 20% oxygen are essentially preadapted or preselected to survive under conditions of oxidative stress (Davies, 1999). Furthermore, the chronic exposure of cell lines to various levels of oxidative stress over several generations is likely to select for pre-existing or mutant phenotypes that confer oxidative stress resistance (Davies, 1999). It appears therefore logical, that the altered trinucleotide repeat profile of cultured cells upon exposure to oxidative stress is mainly based on the selection and rapid growth of mutant cells that exhibited enhanced resistance to oxidative stress, and that happen to carry shorter (in the case of D2763Kc2 cells) or longer repeats (D4132K cells) than most of the other cells in culture.

In addition to a selection event, and independently of an increase or decrease in the expression levels of DNA repair enzymes in oxidative stress resistant cells, as discussed above, the activity levels of the repair proteins, including MMR enzymes, are most certainly altered under conditions of oxidative stress. It may be therefore speculated that altered MMR activity will be involved in the alternative processing of mutation intermediates, contributing to a modified rate of trinucleotide repeat expansion. Non-cytotoxic levels of hydrogen peroxide inactivate both single-base mismatch and insertion/deletion loop repair activities of the MMR system in human cell lines (Chang *et al.*, 2002). This inactivation is most likely caused by oxidative damage of MSH2-MSH6, MSH2-MSH3 and MLH1-PMS2 heterodimers. Reduced MMR activity in *Dmt-D* cells, exposed to high levels of oxidative stress, could account for the altered dynamics of the transgenic repeats in culture, given the involvement of MMR genes as key genetic modifiers of trinucleotide repeat metabolism (Manley *et al.*, 1999b; van Den Broek *et al.*, 2002; Chapter 8).

Moreover, MSH2-MSH6 heterodimers have also been reported to play a direct role in the repair of 8-oxoG•A mispairs and 8-oxoG•C base pairs in *S. cerevisiae*, suppressing the rate of G•C→T•A transversions in *ogg1* mutants (Ni *et al.*, 1999). Thus, it appears that in eukaryotic cells, an MSH2-MSH6-dependent mismatch repair pathway is also involved in the repair of 8-oxoG lesions. The repair activity of MSH2-MSH6 heterodimers on oxidised bases might only be significant under conditions of high oxidative stress, or when OGG1 or possibly other repair proteins are unable to repair the damaged base pair. Therefore, an MSH2-MSH6-dependent pathway might function as a mechanism for detection of otherwise unrepairable levels of oxidative damage. The recruitment of MSH2-MSH6 by oxidative damage may be intimately associated with a decrease of trinucleotide repeat instability (van Den Broek *et al.*, 2002; a more extended discussion on the dynamics of MMR proteins and triplet repeat mutation is presented in Chapter 9).

Alternatively, an excessive amount of hydrogen peroxide may saturate the overall activity of the MMR system, which might be shifted towards the repair of oxidative damage. As a result, significantly lower levels of MMR proteins could be available to act on trinucleotide repeat substrates, and mediate triplet repeat expansion, leading to a change in repeat mutation profile (Manley *et al.*, 1999b; van Den Broek *et al.*, 2002; Chapter 8).

The marked accumulation of shorter repeat tracts in D2763Kc2 cells exposed to hydrogen peroxide may result from multiple deletion events. Hydroxyl radical attack on the deoxyribose in DNA induces strand breaks, which might be expected to contribute to deletion events and structural chromosome rearrangements (Marnett, 2000). When mouse kidney cells, with an *adenine phosphoribosyltransferase* gene heterozygous deficiency, were exposed to hydrogen peroxide, they showed loss of heterozygosity at nonadjacent loci on the same chromosome (Turker *et al.*, 1999). One possibility is that oxidative stress produces a markedly elevated rate of recombination, given the great number of strand breaks induced by hydrogen peroxide. Alternatively, multiple deletions may explain the discontinuous loss of heterozygosity.

Trinucleotide repeats carried by *Dmt-E* kidney cells did not exhibit increases in the mutation rate upon hydrogen peroxide treatment. Similarly, triplet repeat dynamics in D2763E eye cells was not significantly affected by hydrogen peroxide. Several hypotheses may explain these observations. First, the influence of both *cis* and tissue-specific factors on trinucleotide repeat instability may be greater than the effect of environmental agents, thus the lack of an effect of hydrogen peroxide exposure on the triplet repeat dynamics on cell lines that exhibit naturally stable transgenic sequences. Second, and more likely, hydrogen peroxide may still induce trinucleotide repeat instability in these cell lines. However, given the extremely low intrinsic levels of somatic mosaicism detected under standard conditions for both cultures, the enhanced trinucleotide repeat mutation rate would still be too low to be detectable by SP-PCR procedures. Third, and finally, the effects observed *in vitro* with other cell lines, particularly the D2762Kc2 kidney clonal cultures, may simply result from a selection process, which necessarily requires repeat size variability prior to the establishment of the treatment. In the absence of repeat size variability within a particular culture, selection may still occur but will not be detectable by techniques designed to monitor repeat size heterogeneity.

Cultures treated with ethidium bromide did not show any evidence of clonal expansion driven by selection. Nonetheless, a significant decrease in the expansion rate was still observed upon exposure to this chemical, raising serious questions about the induction of trinucleotide repeat instability by ROS. However, the consequences of the ethidium bromide treatment may derive from a wide range of effects that this chemical may inflict upon the metabolism of cultured cells, other than increased levels of oxidative stress. Ethidium bromide interacts directly with DNA, and it is able to alter the migration of putative alternative DNA conformations through agarose gels. The inherent sequence-dependent structural features of DNA such as handedness, planarity and twisting of curved segments can be dramatically changed by ethidium bromide, on the basis of a close relation between sequence-dependent DNA shape and ethidium bromide-induced untwisting of DNA (Brukner *et al.*, 1997). Ethidium bromide not only untwists DNA, but also destabilises the

fine stacking interactions that are important for DNA bending. Under stress-free conditions (absence of ethidium bromide), bent DNA molecules have shown marked gel retardation, relative to untwisting conditions, which were introduced by the addition of ethidium bromide to the gel and running buffer (Brukner *et al.*, 1997). Since trinucleotide repeat sequences are preferential sites for the accumulation of supercoiling and twisting (Gellibolian *et al.*, 1997), it may be possible for ethidium bromide to modify the physical properties of trinucleotide repetitive tracts. Indeed, additional slowly migrating DNA species were clearly observed following Southern blot hybridisation analysis of CAG•CTG-containing PCR products, resolved through agarose gels in the absence of ethidium bromide. Moreover, a slight increase in the electrophoretic mobility of mouse genomic DNA fragments, containing the *Dmt162* transgene, through agarose gels in the presence of ethidium bromide was detected. However, no additional low mobility bands were present when both mouse genomic DNA and bacterial plasmid DNA fragments, both containing long CAG•CTG repeat units, were resolved through agarose gels, with or without ethidium bromide, possibly indicating that the alternative structures clearly observed in the PCR products, are either not present *in vivo* or exist in extremely low levels within the cell (either bacterial or mammalian cells), so that they are not detected by this method. Interestingly, ethidium bromide failed to affect the DNA mobility through agarose gels, when PCR products containing CAG•CTG repeats were incubated with the intercalating agent, preceding electrophoresis. Given the opposite directions of migration of DNA and ethidium bromide during the course of electrophoresis, the intercalating dye may be removed from the PCR products, early during migration through agarose gels, and multiple DNA species were consequently detected.

In addition to the direct interaction of ethidium bromide with DNA, this dye can also inhibit topoisomerase II activity (Thielmann *et al.*, 1993), and result in chromatin hypercondensation in human lymphocytes (Belyaev *et al.*, 1999), thereby possibly causing a decrease in trinucleotide repeat instability, by a mechanism similar to the one previously discussed in Chapter 4.

Finally, it should be noted that oxygen radicals, like all electrophiles, react with many molecules in the cell and can induce a wide range of responses that are not necessarily dependent on DNA damage. Oxygen radicals are known to affect various signal transduction pathways and transcription factors (Allen and Tresini, 2000; Dalton *et al.*, 1999) and to alter cell cycle kinetics (Leroy *et al.*, 2001; Martin *et al.*, 2000). Some of these effects may be the result of DNA damage-induced signalling, but others may not (Marnett, 2000). It was previously reported that concentrations of hydrogen peroxide ranging from 120 to 150  $\mu\text{M}$  cause a temporary growth arrest of mouse fibroblasts (Wiese *et al.*, 1995). At least 40 gene products are involved in the adaptive response: some genes being up-regulated, some being down-regulated. Most of the genes that participate in the mammalian adaptive response are involved in the antioxidant defence; others are damage-removal or repair enzymes. In addition, proto-oncogenes, interleukins, many protein kinases and phosphatases (Davies, 1999) and also RNA species of unknown function (Crawford *et al.*, 1996) are also involved in the response to oxidative stress. Furthermore, oxidants increase cytosolic calcium ions concentration, which is a key messenger in multiple metabolic pathways,

including apoptosis (Dalton *et al.*, 1999; Kowaltowski *et al.*, 2001; Wong and Cortopassi, 1997). It is apparent from the findings of the studies reported by others, that the effects of specific redox perturbations involve a wide range of metabolic pathways, and although they can be similar among organisms and tissues, some redox-active stimuli are highly cell- and tissue-specific (Allen and Tresini, 2000). Therefore, independently of the hypothetical molecular mechanism linking oxidative stress to trinucleotide repeat dynamics, it is important to note that many cellular pathways are affected by ROS, and therefore it is extremely difficult to pinpoint the exact causes of altered repeat stability following exposure to increased oxidative stress.

The findings presented in this chapter suggest that trinucleotide dynamics in cultured cells can be modified by DNA damage induced by oxygen free radicals. Furthermore, the results clearly illustrate that selection of cells exhibiting enhanced resistance to oxidative stress, may play a crucial role in determining the final repeat mutation profile exhibited by *Dmt-D* cultured cells. In most cases the selected cells appear to carry shorter trinucleotide repeats than the overall population, suggesting that the mutations that confer higher resistance to oxidative stress are usually, but not necessarily always, associated with smaller trinucleotide repeat tracts. However, the observations gathered from the *Dmt-D* cell culture system may not reflect the complex *in vivo* scenario, where there is a lower probability for cell selection, followed by clonal expansion. It would be extremely interesting to correlate the levels of oxidative stress *in vivo*, determined by cell sorting coupled with fluorescent techniques (Jakubowski and Bartosz, 2000), with the degrees of somatic mosaicism trinucleotide repeat dynamics *in vivo* determined by SP-PCR. We must also question whether the undoubted ability of dietary antioxidants to decrease DNA oxidation (Ames *et al.*, 1993) has any biological significance in trinucleotide repeat metabolism. The inclusion of radical scavengers, such as vitamin C, vitamin E or carotenoids, in the diet of *Dmt-D* mice, might have an effect on the stability of the transgene. It would be of great relevance to assess the effect of dietary antioxidants on the mutation rate of trinucleotide sequences. Clinical strategies designed to control the levels of oxidants species may even prove to be useful in the therapy of trinucleotide repeat conditions.

## **6. Investigating the effects of multiple genotoxic agents, affecting cell cycle progression, DNA replication and DNA repair, on the dynamics of expanded triplet repeats**

### **6.1. Introduction**

Based on experimental data collected from simple model organisms, DNA polymerase slippage has been put forward as the most favoured mechanism of trinucleotide repeat expansion (Richards and Sutherland, 1994; Wells *et al.*, 1998). However, at present there is no direct evidence to support the hypothesis that repeat length mutations arise during DNA replication in mammalian cells. In reality, there are no obvious relationships between the estimated rates of cell turnover and the levels of somatic mosaicism, neither in human patients (Anvret *et al.*, 1993; Ashizawa *et al.*, 1993; Hashida *et al.*, 2001; Monckton *et al.*, 1995; Thornton *et al.*, 1994) nor in the mouse models that replicate CAG•CTG somatic mosaicism (Fortune *et al.*, 2000; Kennedy and Shelbourne, 2000; Lia *et al.*, 1998; Seznec *et al.*, 2000). Furthermore, monitoring the dynamics of expanded trinucleotide repeats in homogenous *Dmt-D* cell lines, failed to reveal a correlation between proliferative rates, assessed by population doubling times, and levels of trinucleotide stability detected in culture (Chapter 3). Taken together, these results suggest that the mechanism of expansion is not strictly based on cell division, implying the involvement of additional factors, other than DNA replication, in the control of the mutation rate of trinucleotide repeats in mammalian cells.

All eukaryotes have evolved a plethora of mechanisms to minimise DNA damage and maintain DNA integrity and stability. The genomes of eukaryotic cells are under continuous assault by environmental agents (*e.g.* UV light and reactive biochemicals) as well as the by-products of normal intracellular metabolites (*e.g.* reactive oxygen species intermediates and inaccurately replicated DNA). Whatever the origin, genetic damage threatens cell survival and may lead to organ failure, immunodeficiency, cancer and other pathologies. To ensure that cells pass accurate copies of their genomes to the next generation, evolution has provided the cell cycle machinery with a series of surveillance pathways termed cell cycle checkpoints (Abraham, 2001; Bartek and Lukas, 2001; Elledge, 1996). The overall function of the cell cycle checkpoints is to detect damaged or abnormally structured DNA, and to coordinate cell cycle progression with DNA repair. In a broader context, cell cycle checkpoints can be envisaged as biochemical pathways of signal transduction, which sense various types of genetic lesions and induce a multifaceted cellular response, which activates DNA repair and delays cell cycle progression, allowing time for appropriate repair mechanisms to correct genetic damage before they are passed on to the next generation, hence linking the pace of cell cycle phase transitions to the timely and accurate

completion of prior contingent events (Abraham, 2001; Bartek and Lukas, 2001; Elledge, 1996). In response to DNA damage and obstruction of DNA replication, checkpoints can generate signals that arrest the cell cycle in the G<sub>1</sub> phase, slow down S phase, arrest cells in the G<sub>2</sub> phase and induce the transcription of repair genes (Elledge, 1996). When DNA damage is irreparable, checkpoints also have the ability to eliminate such potentially hazardous cells by permanent cell cycle arrest, or by inducing cell death by apoptosis (Bartek and Lukas, 2001; Elledge, 1996; Smits and Medema, 2001). Checkpoint malfunction leads to accumulation of mutations and chromosomal aberrations, which in turn increase the probability of developmental malformations or genetic syndromes and diseases, including cancer (Bartek and Lukas, 2001).

Diverse cell types and cell cycle phases are assumed to share the upstream elements of the checkpoint cascade, responsible for monitoring DNA integrity. In contrast, the downstream checkpoint effectors and their final targets within the cell cycle machinery may differ in G<sub>1</sub>, S or G<sub>2</sub>/M phases (Bartek and Lukas, 2001). During the very earliest stages of checkpoint activation, DNA damage sensors relay information, via a still elusive mechanism, to members of a family of phosphoinositide 3-kinase related kinases (PIKKs). In mammalian cells two PIKK family members, ATM (ataxia-telangiectasia mutated) and ATR (ATM and Rad3-related) play critical roles in early signal transmission through cell cycle checkpoints (Abraham, 2001). ATM and ATR respond to DNA damage in fundamentally different ways: ATM becomes catalytically active, whereas ATR redistributes into nuclear foci, where it presumably gains access to its substrates (Abraham, 2001). Both kinases regulate other downstream players, such as p53, being therefore proposed as central players in DNA damage checkpoint (Abraham, 2001; Bartek and Lukas, 2001; Smits and Medema, 2001).

Two conditions allow mitosis to occur without repair of DNA damage or completion of DNA synthesis: mutational inactivation of the checkpoint control and exposure to caffeine or a caffeine-like agent (Lau and Pardee, 1982; Schlegel and Pardee, 1986). Caffeine was first reported to potentiate the lethality of the alkylating agent nitrogen mustard, 2-chloro-*N*-(2-chloroethyl)-*N*-methylethanamine, on baby hamster kidney cells, by inducing cells to undergo mitosis before properly repairing lesions in their DNA (Lau and Pardee, 1982). In the absence of caffeine, cells were arrested in G<sub>2</sub> shortly after treatment with nitrogen mustard. Unscheduled DNA synthesis was detected during the G<sub>2</sub> delay. After an arrest of six hours, when DNA repair of the lesions induced by nitrogen mustard was completed, treated cells proceeded into mitosis and from then on behaved like normal untreated cells. Exposure to caffeine prevented the G<sub>2</sub> arrest, inducing cells to undergo mitosis, forcing them to divide without providing enough time for damaged cells to repair their DNA (Lau and Pardee, 1982; Schlegel and Pardee, 1986). Caffeine abolishes the mammalian G<sub>2</sub> checkpoint through the inhibition of ATM and ATR kinase, thereby inhibiting the phosphorylation of downstream targets in the biochemical cascade that mediates the G<sub>2</sub>/M arrest (Blasina *et al.*, 1999; Zhou *et al.*, 2000). Indeed, caffeine has also been reported to mediate override of checkpoint control in *Schizosaccharomyces pombe*, thereby reducing cell viability (Rowley and Zhang, 1999; Wang *et al.*, 1999b), through the inhibition of Rad3 kinase, the fission yeast kinase related to the

human ATM checkpoint protein (Moser *et al.*, 2000). In conclusion, caffeine acts by modifying the function of cell cycle progression control which sensors DNA damage.

Ultraviolet (UV) light is a potent mutagenic and genotoxic agent, capable of causing serious damage, which can ultimately result in mutagenesis and cell killing. In contrast to caffeine, UV-induced DNA damage may lead to an arrest in cell cycle progression (Franchitto *et al.*, 1998), increased transcription of genes involved in DNA repair and DNA replication (Weinert, 1998). UV light is absorbed by nucleic acids and produces several types of DNA damage, which interfere with DNA replication and transcription. Two major classes of mutagenic DNA lesions induced by UV light (UV-C:  $\lambda = 280\text{-}280\text{ nm}$  and UV-B:  $\lambda = 280\text{-}320\text{ nm}$ ) are cyclobutane-pyrimidine dimers (CPDs) and pyrimidine-pyrimidone 6-4 photoproducts (6-4PPs) (Thoma, 1999). After DNA damage by UV radiation, p53 expression is up-regulated, and it can either arrest cell cycle progression, allowing DNA repair, or induce apoptosis (Levine, 1997).

Several mechanisms of repair have been identified for UV-damaged DNA. Nucleotide excision repair (NER) is the major pathway to remove UV-induced DNA lesions from the mammalian genome, hence preventing mutagenesis and cell death (Thoma, 1999). NER in normal human fibroblasts, monitored as unscheduled DNA synthesis, was strongly induced by irradiation with UV-C (Roza *et al.*, 1985), whereas NER-deficient hamster cells were reported to be dramatically more sensitive to UV-C induced apoptosis than proficient cells (Dunkern *et al.*, 2001). Mutations in the *mutS* or *mutL* bacterial mismatch repair (MMR) genes, render the cells moderately sensitive to UV light (Mellon and Champe, 1996). In addition, UV-induced mutations are elevated in *E. coli* strains carrying *mutS*, *mutL* and *mutH* mutations (Liu *et al.*, 2000). The increased sensitivity of yeast strains carrying mutation in both MMR and NER systems, compared with NER single mutants suggested that an MMR-dependent process could serve as an alternate, yet minor, pathway for the repair of UV damage in yeast cells (Bertrand *et al.*, 1998). In addition, mutations in *MSH2* or *PMS2* genes, in cancer cell lines, were found to render human cells slightly more sensitive to UV light (Mellon and Champe, 1996; Mellon *et al.*, 1996). Taken together, these results demonstrate an association between MMR proteins and the NER, and implicate the involvement of MMR enzymes in the repair of UV-induced DNA damage.

Nucleoside antimetabolites constitute one of the major groups of anticancer drugs, which are also able to cause serious DNA damage, inhibit replication and repair, alter cell cycle progression and cause apoptosis. As implicit in their name, these compounds owe their antitumour activity to an ability to masquerade as natural intermediates and thereby inhibit critical biochemical pathways, such as DNA or RNA synthesis. Two deoxycytidine analogues, 1- $\beta$ -D-arabinofuranosylcytosine (araC) and 5-aza-2'-deoxycytidine are the drugs of choice in the treatment of acute myeloid leukaemias (Grant, 1998; Hatse *et al.*, 1999).

AraC is efficiently phosphorylated by deoxycytidine kinase and incorporated into nuclear DNA in both human T-lymphoblastoid and Chinese hamster ovary cells (Zhu *et al.*, 2000). Inhibition of DNA synthesis follows the incorporation of araC into growing DNA strands, whereupon the incorporated nucleotide analogue acts as a potent block to further chain elongation by purified DNA polymerases  $\alpha$ ,  $\beta$  and  $\epsilon$  (Harrington and Perrino, 1995; Mikita and Beardsley,

1988; Ohno *et al.*, 1988; Parker and Cheng, 1987; Perrino and Mekosh, 1992; Townsend and Cheng, 1987). At concentrations that inhibit DNA synthesis, araC induces a dramatic accumulation of Okazaki fragments in leukaemia cells, implying that DNA polymerase elongation is seriously compromised *in vivo* by the incorporation of araC in the nascent DNA strand (Catapano *et al.*, 1991). However, despite being an inhibitor of DNA elongation, the incorporation of araC into newly synthesised strands does not terminate the production of full-length DNA, instead it is believed to cause a profound decrease in the rate of DNA synthesis and therefore replication (Wills *et al.*, 1996; Wills *et al.*, 2000). More interestingly, araC is responsible for a specific alteration in replication fork progression, affecting DNA synthesis on the two strands of the replication fork to a different extent, and causing an imbalance between leading and lagging strand synthesis. Analysis of newly replicated DNA, synthesised in the presence of this antimetabolite, revealed a greater inhibition of leading strand synthesis, compared with lagging strand synthesis (Carbone *et al.*, 2001). In addition to blocking semiconservative DNA synthesis, araC has also been reported to inhibit the polymerisation step during excision repair pathways (Fram and Kufe, 1985).

Although 5-azacytidine is another efficient anti-cancer drug, it shows high cytotoxicity at concentrations that do not inhibit DNA synthesis (Momparler and Goodman, 1977). When incorporated into DNA, this antimetabolite causes extensive hypomethylation in dividing cells. This is due to the covalent binding of DNA methyltransferase to 5-azacytidine incorporated into DNA (Juttermann *et al.*, 1994; Santi *et al.*, 1984). Treatment of mammalian cell lines with 5-azacytidine or its deoxyribose congener, 5-aza-2'-deoxycytidine, has resulted in a variety of altered phenotypes, including changes in chromosome structure, gene expression and cellular morphology (Parrow *et al.*, 1989; Singal *et al.*, 1997), and in the induction of apoptosis (Canova *et al.*, 1998; Kajikawa *et al.*, 1998; Kizaki *et al.*, 1992).

DNA hypomethylation, induced by 5-azacytidine, may also have great impact on DNA repair, by affecting the strand discrimination step by MMR enzymes. In *E. coli*, adenine hemimethylation within GATC sequences provides the information that allows for strand discrimination. MutS, MutL and MutH act in coordinated fashion to make a single-strand nick at a hemimethylated GATC site in the newly replicated strand (Buermeyer *et al.*, 1999). Given the similarities between both prokaryotic and eukaryotic MMR systems, there is at least the possibility that the process might also be methyl group-directed in eukaryotes (Bellacosa, 2001; Cleaver, 1994). Alternatively, it has been proposed that mammalian strand discrimination relies on DNA termini generated at the replication fork, which may provide the strand-targeting signal *in vivo*. The proliferating-cell nuclear antigen (PCNA) might facilitate this process, by physically linking the replicative DNA polymerase with MMR proteins (Umar *et al.*, 1996). Regardless of the nature of the strand discrimination signal in mammalian cells, it has been reported that promoter hypermethylation of MMR genes (Herman *et al.*, 1998; Veigl *et al.*, 1998) and mutations in methyl-CpG binding proteins are associated with microsatellite instability in humans (Bader *et al.*, 1999; Riccio *et al.*, 1999).

Nonsteroidal anti-inflammatory drugs (NSAIDs) are well known cancer preventatives, which has been largely attributed to their antiproliferative and pro-apoptotic activities. Aspirin, in

particular, has been reported to display profound antiproliferative effects on tumour cell lines, alter the cell distribution through the cell cycle phases and induce apoptosis (Arango *et al.*, 2001; Li *et al.*, 2000b; Ricchi *et al.*, 1997; Sheng *et al.*, 1997; Shiff *et al.*, 1995; Stark *et al.*, 2001). NSAID treatment of colon tumour cells results in a dramatic increase in arachidonic acid, which in turn stimulates the conversion of sphingomyelin to ceramide, a mediator of apoptosis (Chan *et al.*, 1998). Several epidemiological studies have shown that prolonged use of aspirin is associated with reduced risk of colorectal cancer by as much as 40-50% and it also appears effective in other gastrointestinal cancers, such as esophageal and gastric carcinoma as well as several other tumour types (Coogan *et al.*, 1999; Cramer *et al.*, 1998; Egan *et al.*, 1996; Funkhouser and Sharp, 1995; Garcia-Rodriguez and Huerta-Alvarez, 2001; Schreinemachers and Everson, 1994; Suh *et al.*, 1993; Thun *et al.*, 1993). The best-known mechanism of NSAIDs is the inhibition of cyclooxygenases, the enzymes that convert arachidonic acid to prostanoids in the inflammatory response (DuBois and Smalley, 1996; Gustafson-Svard *et al.*, 1997; Lu *et al.*, 1995; Xu *et al.*, 1999). Interestingly, it has been shown that tumours contain high-level expression of cyclooxygenase-2, which appears to inhibit apoptosis in colon carcinogenesis (Patrignani, 2000; Sano *et al.*, 1995; Tsujii and DuBois, 1995). Western blot and immunocytochemical analyses of NSAID treated cell populations suggested that some pro-apoptotic proteins were markedly induced upon exposure to aspirin (Ruschoff *et al.*, 1998). However, in contrast with araC and 5-azacytidine, aspirin was reported to reduce genomic instability. In particular, aspirin was reported to stabilise dinucleotide sequences in human cells carrying *MLH1*, *MSH2* or *MSH6* mutations (Ruschoff *et al.*, 1998). The stabilisation of the dinucleotide tracts was confined to non-apoptotic cells, while higher levels of microsatellite instability were still detected in apoptotic cells, which were eliminated from the growing population. It was concluded that aspirin reduces microsatellite instability by means of genetic selection (Ruschoff *et al.*, 1998).

Since the genotoxic agents described above, affect distinct aspects of the cellular metabolism, including cell cycle progression, DNA replication and repair, and induction of programmed cell death, we sought to investigate their effects on the dynamics of expanded CAG•CTG trinucleotide repeats carried by *Dmt-D* cultured cells, aiming to gain greater insight into the biology of triplet repeats, and in particular, into the possible modifiers of the mutation mechanism.

## 6.2. Results

### **6.2.1. Association between increased sensitivity to UV radiation and greater levels of trinucleotide repeat instability**

If the mechanism of trinucleotide repeat mutation were strictly dependent on DNA replication and cell division, a tight correlation between the levels of repeat size variability and the rates of cell turnover in culture should be expected for the different cell lines established from *Dmt-*

D mice. Nevertheless, the results presented in Chapter 3 reject such a simple association and suggest that factors, other than DNA replication, must be involved in the mechanism of trinucleotide repeat mutation. Being involved in the maintenance of genomic integrity, the multiple DNA repair systems might be considered as potential modifiers of repeat metabolism *in vivo*. To test this hypothesis, the sensitivity of two *Dmt-D* cell lines to UV-C radiation ( $\lambda = 254$  nm) was assessed. Assuming that greater survival to UV exposure is associated with higher DNA repair activity of UV-induced damage during cell cycle arrest (Weinert, 1998), the sensitivity of a cell line to UV damage might be considered as an approximate indicator of the efficiency of its DNA damage checkpoints, repair pathways, and/or efficiency of the MMR signalling cascade leading to apoptosis. The cell lines selected to perform this study were both derived from the same animal, yet from different tissues, exhibiting distinct levels of repeat length variability in culture: D2763L lung cells show low levels of repeat instability, whereas D2763K kidney cells exhibit greater levels of repeat size heterogeneity (Figure 3.4). Despite the differing degrees of repeat instability, both cell lines proliferated at identical rates, with a population doubling time of  $\sim 30$  hours (Table 3.2).

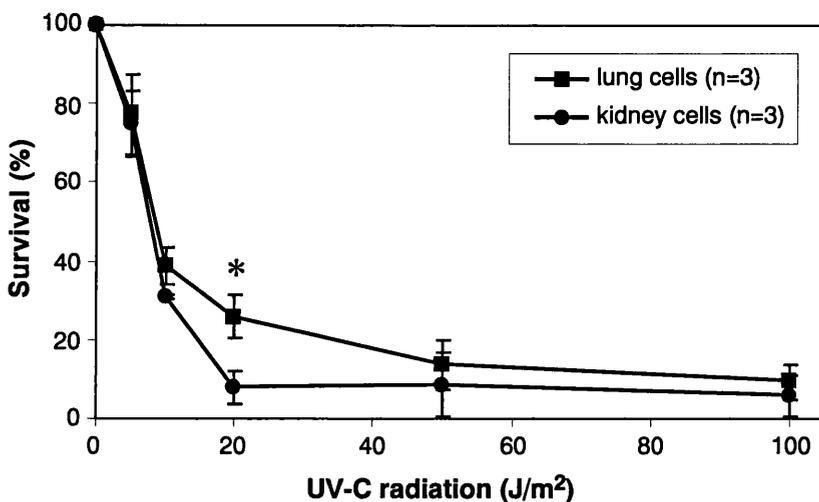
Three replicates of each cell line were exposed to increasing doses of UV-C light, ranging from 10 to 100 J/m<sup>2</sup>, allowed to recover for a period of time equivalent to one population doubling ( $\sim 30$  hours), to ensure that all cells underwent at least one cell cycle during this period. At the end of this period, cell survival was determined by the acridine orange and ethidium bromide method (Section 2.4.6). Both cell lines exhibited a rapid increase in the rate of cell death with greater doses of UV-C radiation (Figure 6.1). Kidney cells (D2763K), however, showed greater levels of cell death compared to lung cells (D2763L). A significantly  $\sim 3$ -fold higher cellular survival was detected in lung cells following an exposure to 20 J/m<sup>2</sup> UV-C ( $p=0.0106$ , two-tailed *t*-test).

In summary, the results appear to suggest an association between greater levels of trinucleotide repeat instability and enhanced sensitivity to UV-C exposure. A similar association between higher levels of repeat trinucleotide mutation in *Dmt-D* cultures and increased sensitivity to hydrogen peroxide was previously described for the same cell lines (Section 5.2.1).

### **6.2.2. Effects of exposure to UV radiation on the dynamics of expanded trinucleotide repeats**

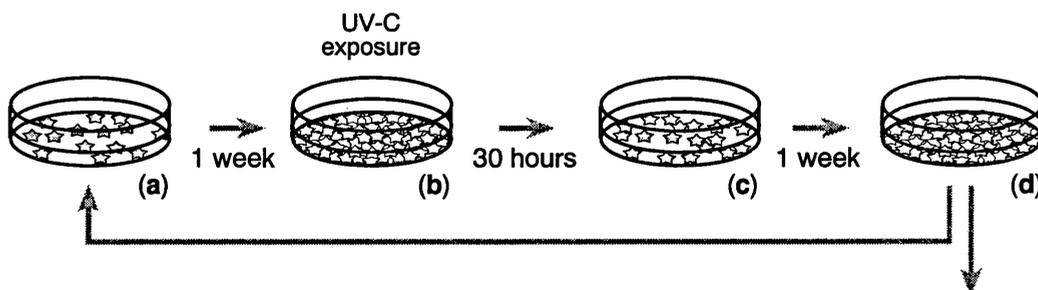
To test the involvement of the DNA repair systems in the molecular mechanisms driving trinucleotide repeat instability, DNA repair activity was induced by the exposure of *Dmt-D* cultured cells to UV-C radiation (Figure 6.2). Confluent cultures were irradiated with UV-C light ( $\lambda = 254$  nm), resulting in a variable degree of cell death depending on the dose of UV-C light used. Once the cells reached confluency they were passaged as usual, and the cells plated on a fresh tissue culture dish were grown until they became 80-100% confluent, to be then subjected to a second UV-C exposure. The remaining cells, which were not seeded, were used to collect DNA and size the transgenic CAG•CTG repeats by sensitive SP-PCR techniques (Figure 6.2).

The effects of the exposure to UV-C radiation on the dynamics of expanded CAG•CTG trinucleotide repeats were initially investigated in two independent cell lines derived from the same



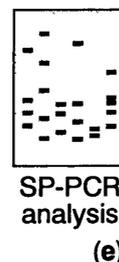
**Figure 6.1. Sensitivity of *Dmt-D* cultured cells to UV-C light exposure.**

Three replicate *Dmt-D* lung and kidney cell cultures, derived from the same mouse, were exposed to increasing doses of UV-C radiation ( $\lambda = 254 \text{ nm}$ ), ranging from 5 to 100 J/m<sup>2</sup>. The cells were initially washed in 1X PBS twice. The washing buffer was removed and the cells exposed to UV-C light in a Stratagene® UV crosslinker 2400. Fresh growth medium was added to the cultures and the cells were allowed to recover for 30 hours. The cell viability was measured by the acridine orange and ethidium bromide method and the percentage of survival was plotted as a function of the dose of UV-C light the cells were exposed to. The graph shows a decrease in cell survival with increasing doses of UV-C radiation. D2763K kidney cells exhibited slightly greater sensitivity to UV-C light than D2763L lung cell cultures. The difference between the sensitivity of the two cell lines to UV-C radiation was statistically significant following an exposure to 20 J/m<sup>2</sup> (\*,  $p=0.0106$ , two-tailed *t*-test).



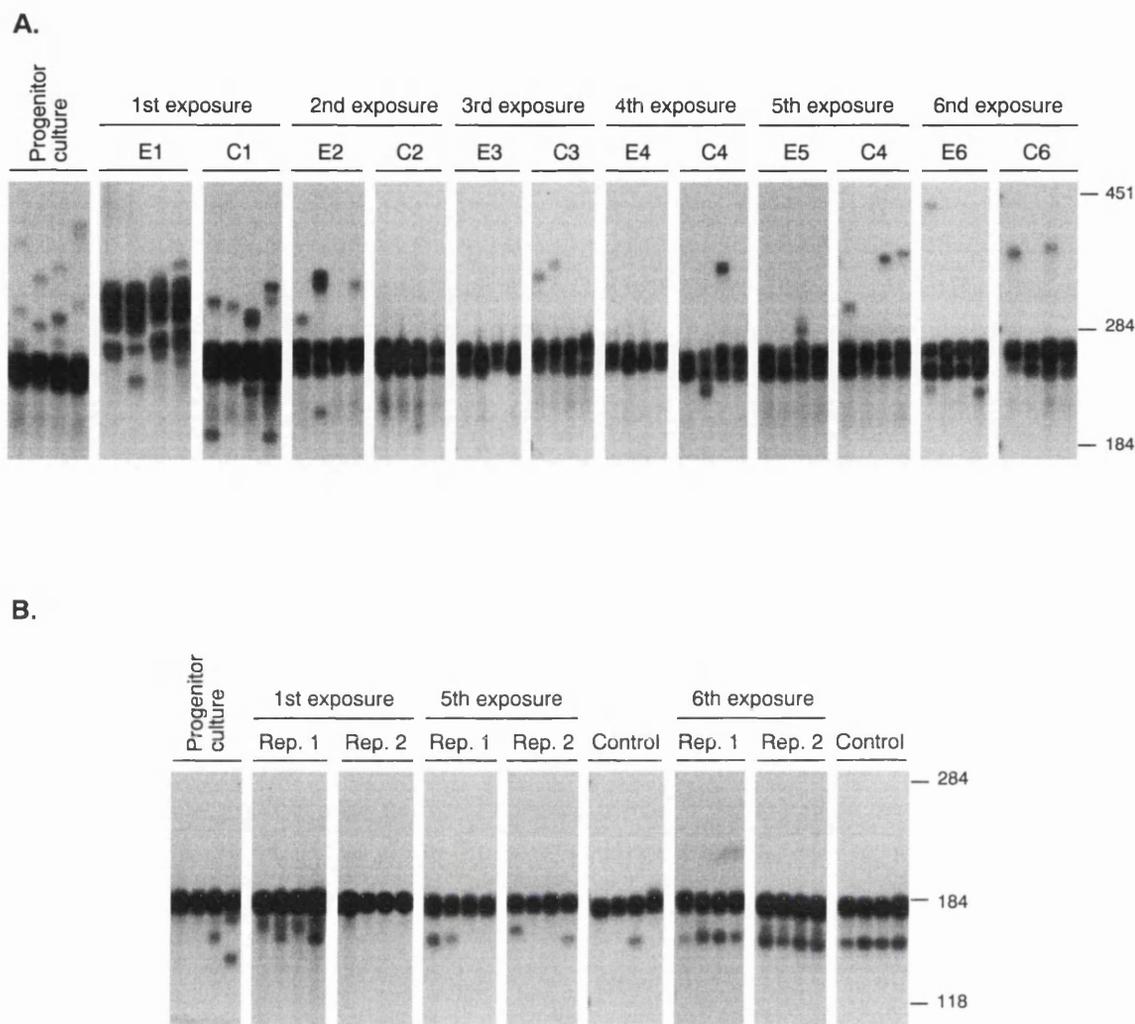
**Figure 6.2. Experimental design to assess the effect of UV-C light exposure on trinucleotide repeat dynamics in *Dmt-D* cultured cells.**

*Dmt-D* cell cultures (a) were grown to 80-100% confluency (b) in standard growth medium. The cells were then washed twice with 1X PBS prior to exposure to UV light. The washing buffer was removed and the cells exposed to 5, 10 or 20 J/m<sup>2</sup> UV-C radiation ( $\lambda = 254 \text{ nm}$ ) in a Stratagene® UV crosslinker 2400 (b) and allowed to recover for a week in standard growth medium, following UV-induced cell death (c). Fresh medium was added to the cells every 2-3 days. Once the cultures became confluent (d), the cells were digested with trypsin, seeded at a ratio of 1:40 and exposed to UV-C radiation when they reached 80-100% confluency once again. The remaining cells were collected and used to extract DNA, and to assess trinucleotide repeat variability by SP-PCR analysis (e).



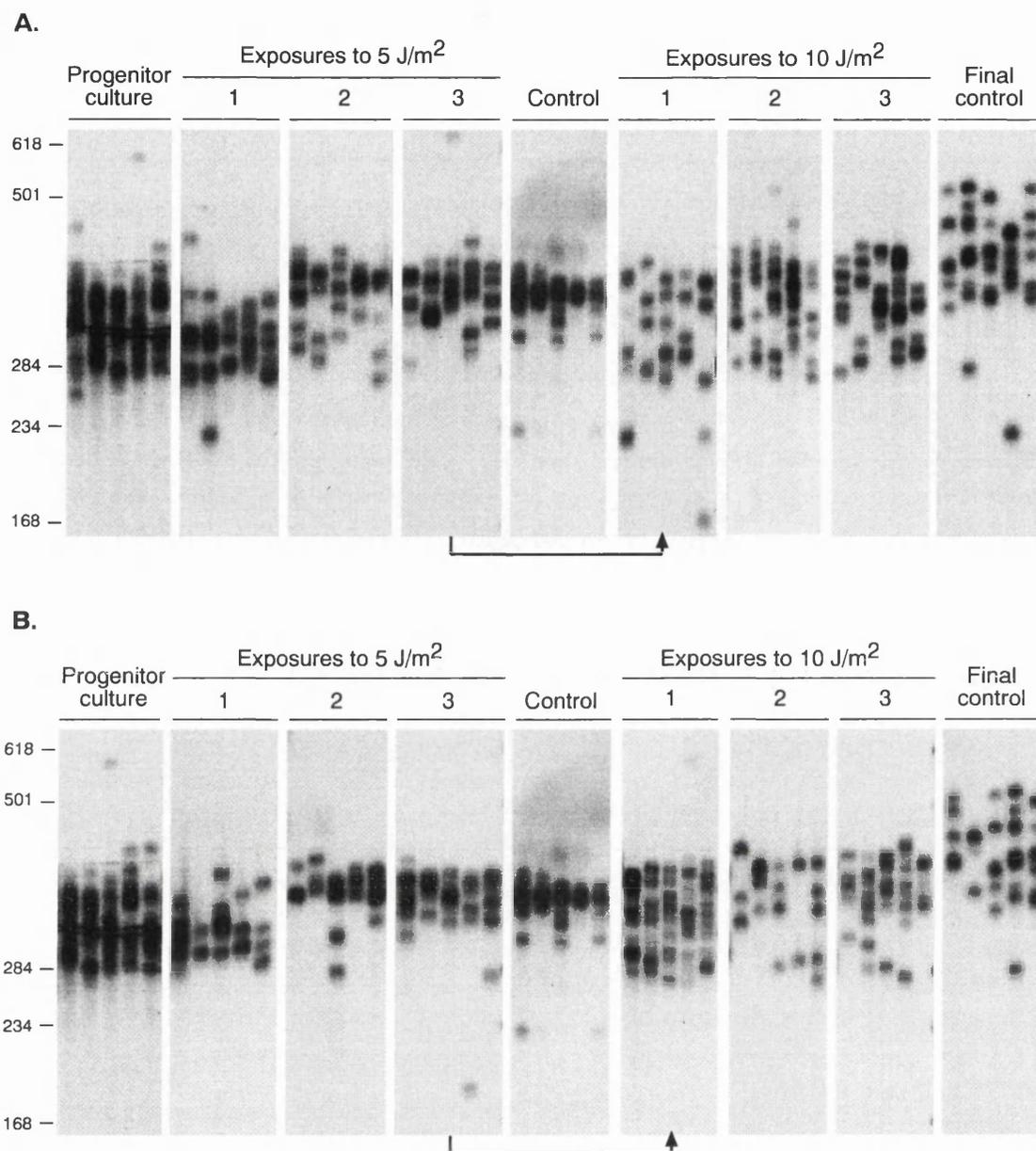
six-month-old mouse: D2763L and D2763K cells (Figure 6.3). Both cell lines were exposed to  $20 \text{ J/m}^2$  UV-C light for six consecutive times, resulting in great levels of cell death (~70-90%), significantly higher for the D2763K cell line. Interestingly, both the levels of repeat size variability and the average repeat size carried by D2763K kidney cells increased dramatically following the first UV-C exposure, compared to the control cells (Figure 6.3.A). Despite the great effect observed following the first exposure of kidney cells to UV light, lower levels of repeat length heterogeneity and lower average repeat numbers, identical to those found in control cultures, were detected following subsequent exposures (Figure 6.3.A). The surprising results obtained with the kidney cell line were not replicated with lung cell cultures. The exposure of D2763L lung cells, which exhibit intrinsically low levels of repeat size instability, failed to cause any obvious change in the repeat number profile detected by high DNA input SP-PCR analysis (Figure 6.3.B).

In an attempt to clarify these inconclusive results, a clonal kidney cell line (D2763Kc2, Figure 3.8.A) that carried rapidly expanding CAG•CTG repeats, was used in a very similar experimental approach. Given that  $20 \text{ J/m}^2$  UV-C is a high dose of radiation, which causes the death of more than 90% of kidney cells in culture (Figure 6.1), lower doses of radiation were used in subsequent assays. Two replicate cultures derived from the same progenitor cells were initially exposed to  $5 \text{ J/m}^2$  UV-C radiation three times as described in Figure 6.2. The repeat size variability was monitored following each exposure by SP-PCR analysis, and the repeat number profile observed after three exposures was compared to a control culture maintained for the same period of time under normal conditions (Figure 6.4). In contrast to the observations described for the D2763K cell line when exposed to  $20 \text{ J/m}^2$ , the exposure of D2763Kc2 kidney cells to  $5 \text{ J/m}^2$  UV-C light appeared to result in a decrease in the average repeat length. However, consistent with the results described previously, the effect observed following a single exposure to  $5 \text{ J/m}^2$  UV-C light was not replicated following subsequent exposures (Figure 6.4). In fact, a dramatic increase in the average repeat number was observed in both replicates following the second exposure to  $5 \text{ J/m}^2$ , and no major differences between the repeat size variability in cells exposed to  $5 \text{ J/m}^2$  UV-C for three times and control cells was detected by high DNA input SP-PCR analysis. An apparent greater level of repeat size heterogeneity might have developed in UV-C exposed cells, compared to the final control culture. This result could have only been confirmed by single molecule analysis. Following three consecutive exposures to  $5 \text{ J/m}^2$  UV-C, the same cells were then exposed to  $10 \text{ J/m}^2$  UV-C for three additional times, as described before, and the degree of trinucleotide repeat instability compared with a single final control (Figure 6.4). In agreement with the results previously described, the exposure of D2763Kc2 kidney cells to  $10 \text{ J/m}^2$  results in a decrease in the average repeat size and a simultaneous increase in the repeat length variability (Figure 6.4). However, similarly to the data presented above, consecutive exposures to  $10 \text{ J/m}^2$  UV-C failed to cause a progressive decrease in the average repeat size, or a continuous increase in the repeat number variability in culture (Figure 6.4). Nevertheless, the repeat profile detected in repeatedly exposed cells by high DNA input SP-PCR analysis, is clearly distinct from that exhibited by control cells; the former being characterised by a much lower average repeat number.



**Figure 6.3. UV-C exposure and expanded CAG·CTG repeat dynamics in D2763K and D2763L cells.**

The autoradiographs show representative SP-PCR analyses of *Dmt* transgenic sequences extracted from kidney (**A**) and lung (**B**) cultured cells. The cells were exposed to 20 J/m<sup>2</sup> of UV-C radiation ( $\lambda = 254$  nm) for six times, as described in Figure 6.2, and the trinucleotide repeat instability assessed by SP-PCR (panels E for D2763K cells). Control cells were maintained under standard growth conditions for the same period of time (panels C for D2763K cells). An average of 10 to 30 transgene molecules from each cell sample were independently amplified in replicate reactions. The molecular weight markers are shown on the right, upon conversion into CTG repeat numbers. (**A**) D2763K kidney cells exhibited greater expansion-biased repeat size variability following the first exposure to UV light, compared to control cells. However, following subsequent exposures, the degree of repeat instability decreased to levels that were indistinguishable from those detected in control cultures. (**B**) Two replicate cultures of D2763L lungs cells were exposed to UV-C light. No effect of UV-C exposure on trinucleotide repeat dynamics was detected.



**Figure 6.4. UV-C exposure and expanded CAG·CTG repeat dynamics in D2763Kc2 kidney cells.**

The autoradiographs show representative SP-PCR analyses of DNA samples extracted from D2763c2 kidney cells. Two replicate cultures (**A** and **B**) derived from the same progenitor culture were exposed to 5 J/m<sup>2</sup> of UV-C radiation ( $\lambda = 254$  nm) for three consecutive times, as described in Figure 6.2, and the trinucleotide repeat instability assessed by SP-PCR techniques and compared with a control culture maintained under standard growth conditions for the same period of time. The same cells were subsequently exposed to 10 J/m<sup>2</sup> for three times, and the trinucleotide repeat dynamics monitored by SP-PCR analysis and compared with a control culture grown for the same period of time under normal conditions (final control). An average of 10 to 30 transgene molecules were amplified in independent reactions. The molecular weight markers, converted into CTG repeat numbers, are shown on the left. The initial exposure to 5 J/m<sup>2</sup> and 10 J/m<sup>2</sup> appear to induce a decrease in the average trinucleotide repeat size in cultured D2763Kc2 kidney cells. However, subsequent exposures failed to replicate the same effect.

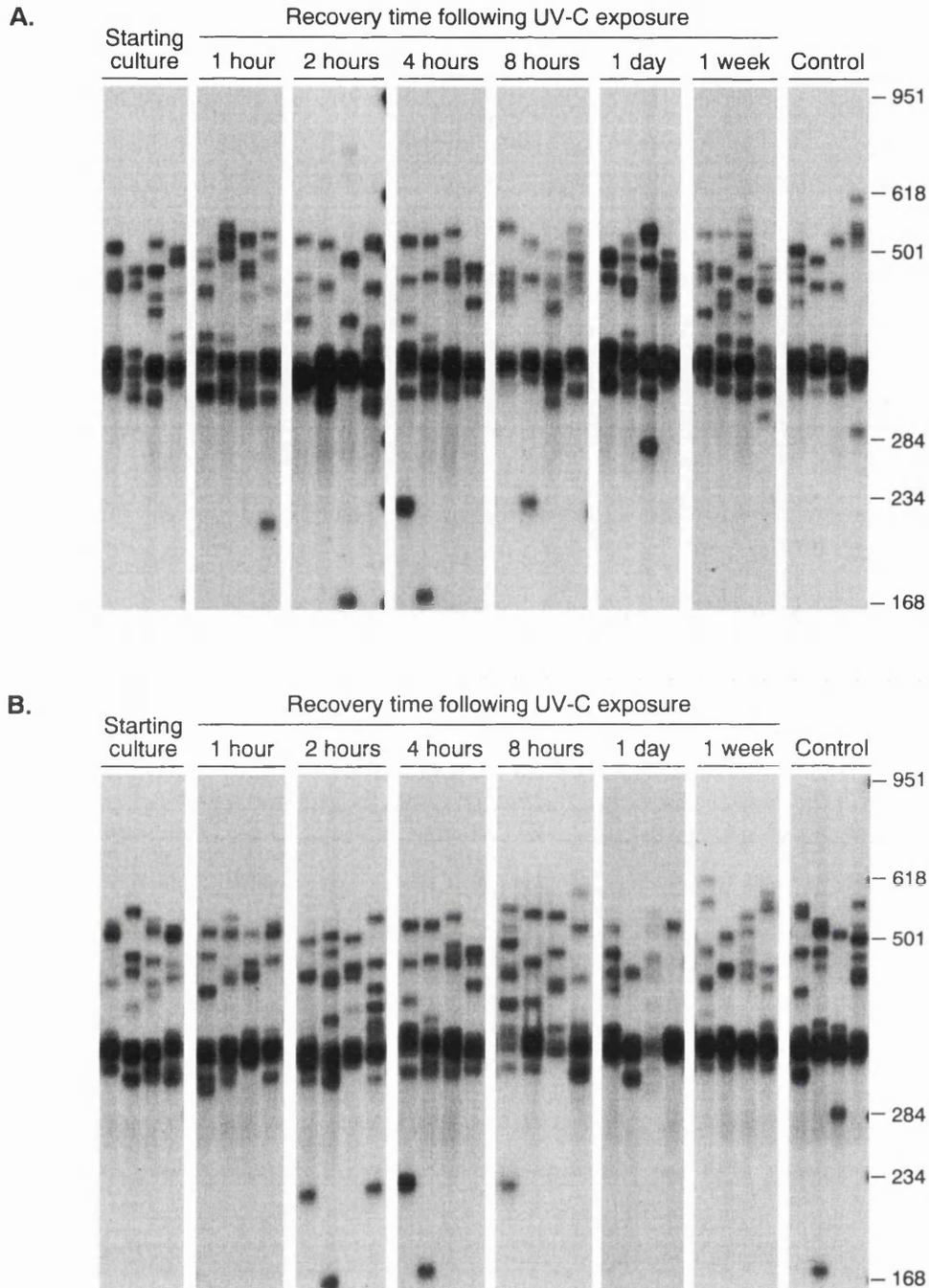
Changes in the mutation profile in culture might be caused by the preferential survival of cells carrying particular repeats sizes, which could be associated with mutations elsewhere in the genome, which not only control cell survival upon stress conditions, but also trinucleotide repeat metabolism. The repeats carried by the most resistant cells may either be longer than the average repeat size, which would give rise to an apparent overall repeat expansion, or shorter, which would result in an apparent overall repeat contraction. This hypothesis does not imply a progressive effect following repeated UV-C exposures, as observed. In this case the change in repeat number would take place in a cell division independent way, and over the first hours following the exposure to UV-C light, independently from the repeat size mutation trend (expansion or deletion). To test this hypothesis two replicate confluent cultures, derived from the same progenitor D2763Kc2 culture, yet at a later passage corresponding to an additional 130 population doublings, were exposed to 10 J/m<sup>2</sup>. The cells were collected after a recovery period ranging from one hour to one week and the repeat size variability assessed by SP-PCR amplification. Unfortunately, the analysis failed to reveal any effect of UV-C exposure on the trinucleotide repeat profiles over time, as detected by high DNA input SP-PCR analysis (Figure 6.5). Ideally, this experiment should have been done with replicates of the D2763Kc2 culture previously studied, collected at an earlier stage, given the likelihood of continuous accumulation of mutations over time, which may interfere with the effects of UV light on trinucleotide metabolism (Figure 6.4).

In summary, UV-C exposure has little effect, if any, on the trinucleotide dynamics of stable expanded CAG•CTG repeats carried by *Dmt-D* lung cells. In contrast, the dynamics of unstable CAG•CTG trinucleotides carried by *Dmt-D* kidney cells is usually, but not always, affected by exposure to UV-C. However, the final trend in the repeat number change may differ between different cultures.

### **6.2.3. Inhibition of DNA damage checkpoint by caffeine induces greater levels of trinucleotide repeat instability**

Caffeine is able to uncouple DNA repair and replication from the progression of the cell cycle (Lau and Pardee, 1982; Schlegel and Pardee, 1986). This is now known to be achieved by directly inhibiting ATM kinase activity (Blasina *et al.*, 1999). Progression through the G<sub>2</sub>/M DNA-damage checkpoint, before replication and repair are completed, inevitably leads to a general increase in the mutation rate. Thus, it is possible that “forced” progression through the cell cycle may also destabilise difficult to replicate expanded trinucleotide repeats. To test this hypothesis, *Dmt-D* cells were exposed to caffeine and the repeat size variability assessed by SP-PCR sensitive techniques at the end of the treatment.

D2763Kc2 kidney cell cultures were continuously exposed to 2 mM of caffeine for 130 days and 80 population doublings, which resulted in an increase of 38% in the population doubling time compared to cells growing under standard conditions, in the absence of the drug (Table 6.1). Trinucleotide repeat size variability was assessed by SP-PCR techniques and compared to time and population doubling control cultures. High DNA input SP-PCR analysis revealed greater expansion



**Figure 6.5. Time course monitoring of the effects of UV-C radiation on the expanded CAG•CTG repeat dynamics in D2763Kc2 cells.**

Two replicate D2763c2 kidney cell cultures (**A** and **B**) were exposed to  $10 \text{ J/m}^2$  UV-C radiation ( $\lambda = 254 \text{ nm}$ ) and the cells collected following a recovery period varying from one hour to one week. Trinucleotide repeat size variability was monitored by SP-PCR analysis, and representative amplifications are shown in the autoradiographs. Two control cultures, consisting of cells grown under standard conditions for one week, were also used for DNA extraction and SP-PCR analysis. An average of 10 to 20 molecules were amplified in independent reactions. The molecular weight markers, converted into CTG repeat number, are shown on the right side. Both replicates failed to reveal an apparent effect of UV-C exposure on the dynamics of transgenic CAG•CTG repeats.

rates in caffeine treated cultures compared to both controls (Figure 6.6.A). To quantify the degree of repeat length variation, a statistical analysis was performed on the repeat numbers determined for a 20 to 50 individual transgene sequences collected from each replicate culture (Figure 6.6.B). The change in repeat number was determined for replicate cultures exposed to caffeine, corrected for both time and cell turnover, and compared to the closest control culture, in terms of time or population doublings, respectively. Caffeine treated cells exhibited a significantly higher median expansion rate compared to time and doubling control cultures (Table 6.1).

**Table 6.1. Summary of the effects of multiple chemical treatments on the dynamics of expanded trinucleotide repeats in D2763Kc2 cells.**

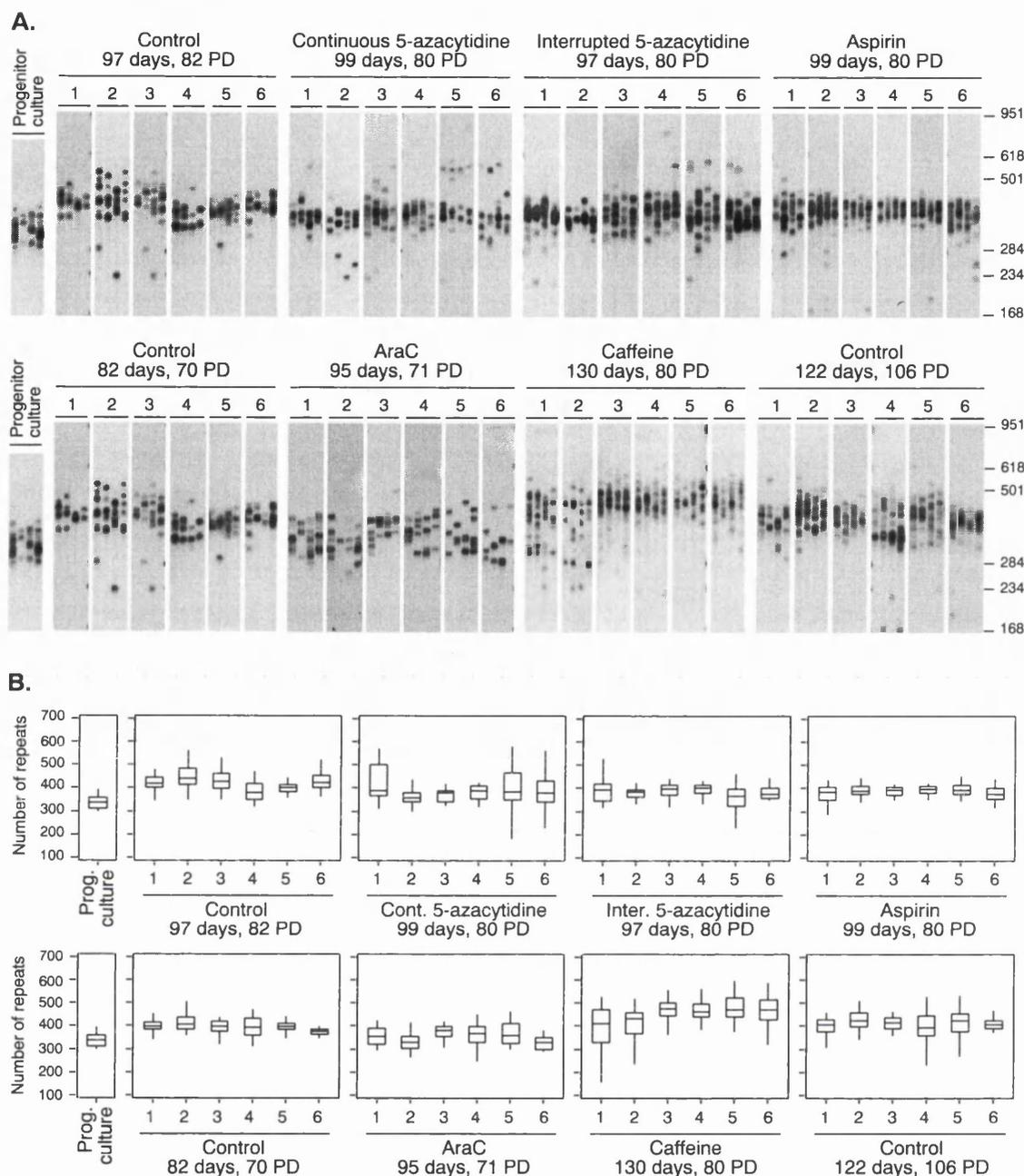
The table shows the time and number of population doublings (PD) of D2763Kc2 cell cultures upon chronic exposure to several chemicals, and the consequent increase in the population doubling time (PDT). The median repeat gain was determined by single molecule SP-PCR analysis, corrected for both time and population doublings, and compared with the closest control cultures in terms of days and PD in culture, respectively (two-tailed Mann-Whitney *U* test).

	Time (days)	PD	Increase In PDT (%)	Median repeat gain per day	Two-tailed Mann-Whitney <i>U</i> test (p)	Median repeat gain per PD	Two-tailed Mann-Whitney <i>U</i> test (p)
Control	82	70	-	0.890	-	1.043	-
Control	97	82	-	0.887	-	1.050	-
Control	122	106	-	0.615	-	0.709	-
Caffeine	130	80	38%	1.006	0.0202	1.634	0.0453
AraC	95	71	13%	0.219	0.0082	0.293	0.0082
Continuous <sup>a</sup> 5-azacytidine	99	80	5%	0.461	0.0202	0.548	0.0202
Interrupted <sup>b</sup> 5-azacytidine	97	80	2.5%	0.547	0.0306	0.663	0.0453
Aspirin	99	80	5%	0.565	0.0453	0.699	0.0453

<sup>a</sup>Continuous 5-azacytidine: cells were treated with 10  $\mu$ M 5-azacytidine for 80 PD.

<sup>b</sup>Interrupted 5-azacytidine: cells were treated with 10  $\mu$ M 5-azacytidine for 40 PD, followed for a further 40 PD in the absence of the chemical.

To further test the effect of caffeine on the dynamics of expanded CAG•CTG repeats, replicate cultures derived from the D4132K cell line were exposed to 2 mM of caffeine for 19 population doublings, over 95 days, resulting in a two-fold higher population doubling time (Table 6.2). The D4132K cell line had only undergone a few population doublings (12 population doublings in 84 days) prior to caffeine treatment, thereby reducing the chances of accumulation of multiple genetic mutations elsewhere in the genome, which may affect the trinucleotide repeat metabolism and mask the effect of caffeine on the expanded CAG•CTG sequence dynamics.



**Figure 6.6. Chemical treatment and expanded CAG-CTG repeat dynamics in D2763Kc2 cells.**

(A) The autoradiographs show representative SP-PCR amplifications of DNA samples extracted from replicate D2763Kc2 cells, cultured for 71-80 population doublings (PD) in the presence of 10  $\mu$ M 5-azacytidine (99 days; replicates labelled "continuous 5-azacytidine"), 5.6  $\mu$ M aspirin (99 days), 500 nM araC (95 days) and 2 mM caffeine (130 days); control cultures maintained for 82 (70 PD), 97 (82 PD) and 122 days (106 PD), and the progenitor culture, from which all cells were derived at day zero. Replicates labelled "interrupted 5-azacytidine" were exposed to 10  $\mu$ M 5-azacytidine for 40 PD, and then grown for a further 40 PD in the absence of the chemical. The molecular weight markers, converted into CTG repeat numbers, are shown on the right. (B) For the quantitative analyses an average of 20 to 50 alleles from each culture were individually sized. The boxplots show the degree of variation observed in treated and control cultures, as previously described (e.g. Figure 4.8). Statistically significant differences ( $p < 0.05$ , two-tailed Mann-Whitney  $U$  test) in the median rates of expansion, corrected for time and population doublings, were observed for all treatments relative to both time and PD controls.

Previous studies indicated that cell turnover is not the most critical factor underlying differences in expansion rates (Chapter 3), therefore the repeat length heterogeneity in treated cultures was compared with the repeat number variability in time control cultures maintained for 86 days and 34 population doublings under normal growth conditions. The analysis revealed that cultures exposed to caffeine carried transgenic repeats that were on average longer than those found on control cells (Figure 6.7.A). The median repeat number for each culture were determined by single molecule analysis performed on 20 to 50 individual transgene sequences (Figure 6.7.B). As described previously for the D2763Kc2 clonal cell line, statistical analysis revealed that caffeine treatment resulted in a significant enhancement of the median repeat number change, relative to the time controls (Table 6.2). Nevertheless, it should be mentioned that in the case of replicate 1 the dramatic increase in repeat length is mainly mediated by the clonal expansion of mutant cells carrying extremely long repeats, which may exhibit enhanced resistance to caffeine (Figure 6.7.A, black arrowheads).

**Table 6.2. Summary of the effects of multiple chemical treatments on the dynamics of expanded trinucleotide repeats in D4132K cells.**

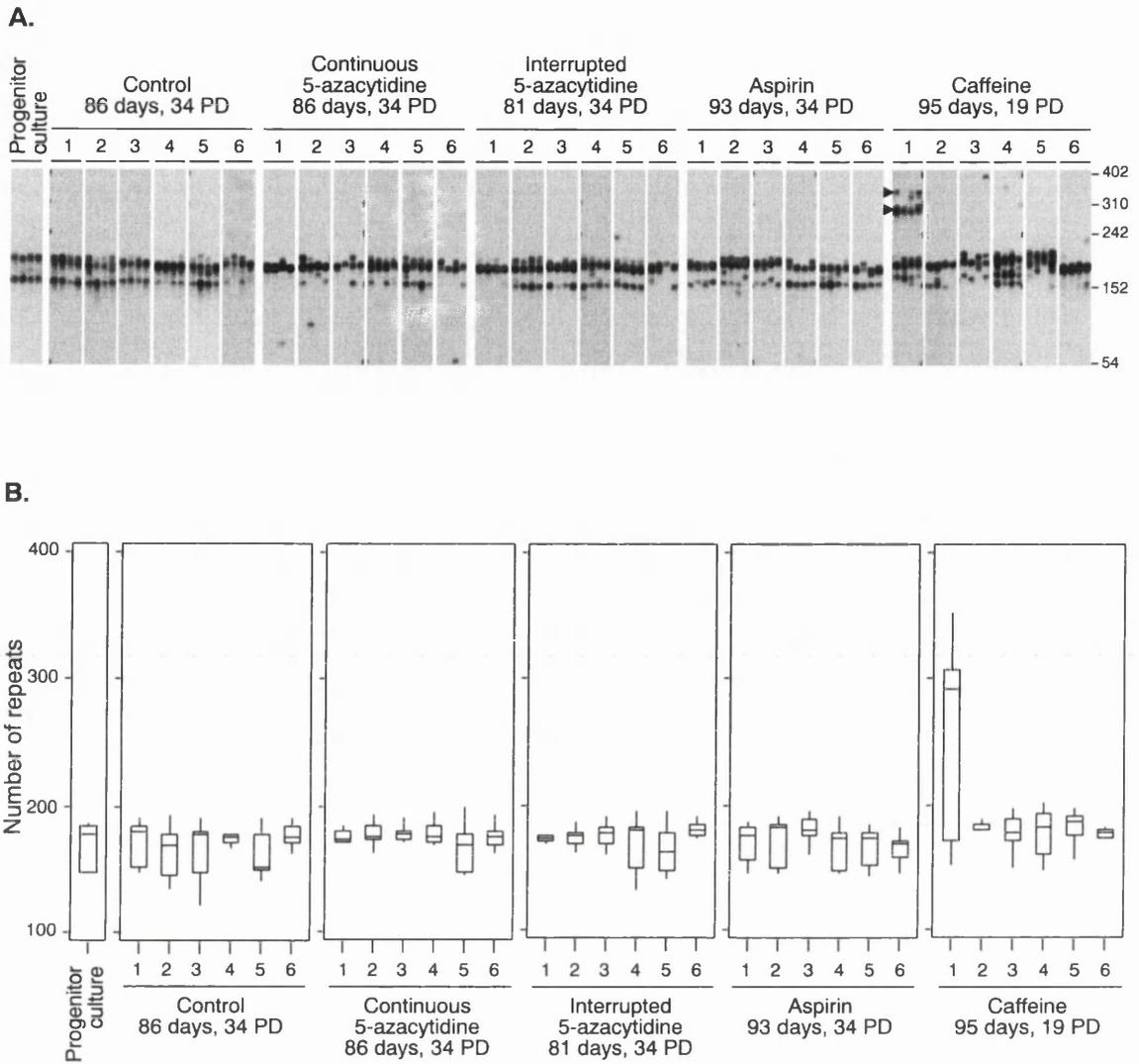
The table shows the time and number of population doublings (PD) of D4132K cell cultures upon chronic exposure to several chemicals, and the consequent change in the population doubling time (PDT). The median repeat change, corrected for days in culture, was determined by single molecule SP-PCR, and compared with the control cultures in terms of days and PD in culture, respectively.

	Time (days)	PD	Change in PDT (%)	Median repeat change per day	Two-tailed Mann-Whitney <i>U</i> test (p)
Control	86	34	-	-0.0279	-
Caffeine	95	19	98%	+0.0313	0.0131
Continuous <sup>a</sup> 5-azacytidine	86	34	0%	-0.0313	0.9362
Interrupted <sup>b</sup> 5-azacytidine	81	34	-6%	-0.0279	0.5752
Aspirin	93	34	8%	-0.0382	0.5752

<sup>a</sup>Continuous 5-azacytidine: cells were treated with 10  $\mu$ M 5-azacytidine for 34 PD.

<sup>b</sup>Interrupted 5-azacytidine: cells were treated with 10  $\mu$ M 5-azacytidine for 17 PD, followed for a further 17 PD in the absence of the chemical.

In conclusion, “forced” progression through the DNA-damage checkpoint induced by exposure to caffeine leads to enhanced levels of trinucleotide repeat instability, consistent with the accumulation of non-repaired DNA-replication slippage errors during mitosis.



**Figure 6.7. Chemical treatment and expanded CAG-CTG repeat dynamics in D4132K cells.** (A) The autoradiographs show representative SP-PCR amplifications of DNA samples extracted from replicate D4132K cells, cultured for 34 population doublings (PD) in the presence of 10  $\mu$ M 5-azacytidine (86 days; replicates labelled “continuous 5-azacytidine”) and 5.6  $\mu$ M aspirin (99 days). Replicates labelled “interrupted 5-azacytidine” were exposed to 10  $\mu$ M 5-azacytidine for 17 PD, and then grown for a further 17 PD in the absence of the chemical. Six replicates were exposed to 2 mM caffeine for 19 PD (95 days). Control cultures maintained for 86 days (34 PD). The progenitor culture, from which all cells were derived at day zero is shown on the left. The molecular weight markers, converted into CTG repeat numbers, are shown on the right. Note evidence of clonal expansion in caffeine treated replicates (black arrowheads,  $\blacktriangleright$ ). (B) For the quantitative analyses, an average of 20 to 50 transgene molecules from each culture were individually sized. The boxplots show the degree of variation observed in treated and control cultures, as previously described (e.g. Figure 4.8). Statistically significant differences ( $p < 0.05$ , two-tailed Mann-Whitney  $U$  test) in the median rates of expansion, corrected for time and population doublings, were observed for caffeine treatment, relative to controls.

#### **6.2.4. Inhibition of DNA polymerase elongation step by araC is associated with lower levels of expansion-biased repeat instability**

In addition to inhibiting ATM kinase activity and thereby bypassing the DNA damage checkpoint, caffeine has also been reported to inhibit replicative DNA synthesis in normal human fibroblasts, transformed hamster cells and human cancer cell lines, in which caffeine is capable of decelerating the progression through S phase (Deplanque *et al.*, 2001; Deplanque *et al.*, 2000). Therefore the effect of caffeine exposure on the stability of trinucleotide repeats may either be caused by the inhibition of cell cycle checkpoint and forced progression into mitosis, or by a more general effect of caffeine on DNA replication. To investigate the effects of DNA replication inhibition on the stability of CAG•CTG repetitive sequences, six replicate *Dmt-D* kidney cultures derived from the D2763Kc2 clonal cell line, were exposed to 500 nM (Table 6.1) and the repeat variability was analysed at the end of the drug treatment, as previously described (Section 6.2.3). The analyses revealed an apparent fall in the average repeat number in araC treated cultures compared to both doublings and time controls (Figure 6.6.A). Single molecule SP-PCR amplification (Figure 6.6.B) and statistical analysis (Table 6.1) revealed a significant and dramatic ~4-fold decrease in the median repeat gain in cell cultures grown in the presence of araC, compared to both controls.

The inhibition of DNA polymerase by araC may therefore cause a significant decrease in the expansion rate of rapidly expanding CAG•CTG unstable repeats.

#### **6.2.5. Induction of DNA hypomethylation by 5-azacytidine and its consequences on the dynamics of CAG•CTG repeats**

Striking correlations between repeat expandability, GC content and proximity to CpG islands have been previously reported in humans (Brock *et al.*, 1999). A potential way in which this effect could be mediated is by DNA methylation at CpG dinucleotides. The nucleotide analogue 5-azacytidine is a potent inhibitor of CpG methylation in mammalian cells and results in global hypomethylation (Juttermann *et al.*, 1994; Santi *et al.*, 1983; Santi *et al.*, 1984). In addition, the DNA methylation status also mediates changes in chromatin structure and may affect DNA-protein interactions, with consequences in gene expression and repair efficiency (Bird and Wolffe, 1999; Razin, 1998).

Given that the *Dmt162* transgene derives from the human *DM1* locus, the CAG•CTG trinucleotide repeats are flanked by a portion of a CpG island, which extends from the last intron of the human *DMPK* gene to the first intron of the downstream *SIX5* gene (Monckton *et al.*, 1997). Adult *Dmt-D* mice exhibit high levels of methylation in the flanking region of the transgenic repeat, which possibility accumulates in an age-dependent manner (G.J. Brock and D.G. Monckton, personal communication). Therefore, exposure of *Dmt-D* mouse cells to 5-azacytidine presents a

valid approach to assess the potential genotoxic effect of this chemical, and more generally, of DNA methylation as a modifier of trinucleotide repeat dynamics.

In order to test the effect of global genomic hypomethylation on the stability of expanded CAG•CTG repetitive sequences, six replicate D2763Kc2 cell cultures were exposed to 10  $\mu$ M 5-azacytidine (Table 6.1) and their repeat size variability assessed by SP-PCR analysis and compared with the repeat profile in six control cultures, as previously described (Section 6.2.3). High DNA input SP-PCR analysis appeared to indicate an apparent decrease in the average size of the transgenic repeat tracts in cells exposed to 5-azacytidine (Figure 6.6.A), despite the accumulation of a few cells carrying very long repeats (~550 CTG units) in some of the replicates (particularly replicates 5 and 6, Figure 6.6.A). Single molecule analysis was carried out in order to quantify the effect of 5-azacytidine treatment on the mutation rate of CAG•CTG repeats in this cell line (Figure 6.6.B), and the expansion rates determined and corrected for both time and population doublings (Table 6.1). Statistical analysis also revealed significantly lower expansion rates in 5-azacytidine exposed cultures, when compared with time and doubling controls (Table 6.1).

The adducts formed between DNA methyltransferase and genomic DNA with 5-azacytidine substitution can sterically inhibit DNA replication, transcription and DNA repair, and may induce mutagenesis in mammalian cells (Amacher and Turner, 1987; Bender *et al.*, 1998; McGregor *et al.*, 1989). 5-Azacytidine has also the ability to disturb chromatin condensation in treated cells mainly in heterochromatic regions (Takahashi-Hyodo *et al.*, 1999). Since 5-azacytidine is randomly incorporated into genomic DNA during replication, it may also be incorporated within the CAG•CTG transgenic sequence. Therefore, the possibility exists that 5-azacytidine may interfere with the physical properties of the repetitive DNA tract, possibly affecting alternative non-B-DNA conformations, which may be formed within the trinucleotide sequence, and consequently the trinucleotide repeat dynamics (Chapter 4). A second approach was consequently followed to determine the effects of genomic hypomethylation caused by exposure to 5-azacytidine on repeat instability, minimising the structural consequences of 5-azacytidine treatment on chromatin structure. D2763Kc2 cells were initially cultured in the presence of 10  $\mu$ M 5-azacytidine for 40 population doublings. Following this initial exposure to the drug, which should have caused extensive genomic hypomethylation, the same cultures previously analysed, were grown for a further 40 population doublings in standard growth medium, in the absence of the cytidine analogue (Table 6.1, "Interrupted 5-azacytidine"). During this period cytosine nucleotides should replace unmethylated 5-azacytidine previously incorporated into the genomic DNA and remain unmethylated. Following the last 40 population doublings in culture, DNA chromatin structure should no longer be affected by the formation of adducts between 5-azacytidine and methyltransferase. Instead, physical properties of DNA should now only depend on its methylation status, rather than on the incorporation of toxic nucleotide analogues.

The repeat size variability in culture was assessed following this short-term treatment, and compared to control cultures, grown in the absence of 5-azacytidine for 80 consecutive population doublings (Figure 6.6.A). High DNA input SP-PCR showed an apparent decrease in average repeat size in replicate D2763Kc2 cultures, temporarily exposed to 5-azacytidine. This result was

confirmed by single molecule analysis (Figure 6.6.B), and proved statistically significant (Table 6.1).

When the expansion rates were compared between D2763Kc2 cultures exposed to 5-azacytidine for 80 population doublings, to those determined for cultures temporarily treated with the chemical for the first 40 population doublings, the analysis failed to reveal statistically significant differences ( $p=0.3785$ , two-tailed Mann-Whitney  $U$  test). This result supports the hypothesis that the lower expansion rates observed in *Dmt-D* cells are correlated to the DNA hypomethylation status, rather than the incorporation of a nucleotide analogue into the genomic DNA *per se*. However, no comparison was performed between the final mutation profiles and the intermediate time point, at which 5-azacytidine was removed from the medium. This analysis should give some insight into trinucleotide repeat dynamics over the last 40 population doublings in culture, in the absence of 5-azacytidine.

As previously described for caffeine treatments, the effects of genomic hypomethylation on trinucleotide repeat metabolism were also studied in D4132K kidney cells. Six D4132K replicate cultures were continuously or temporarily exposed to 5-azacytidine (Table 6.2) and the repeat size variability in culture analysed and quantified as previously described (Figures 6.7). Despite the higher median rate of repeat contraction exhibited by D4132K cells exposed to 5-azacytidine, either continuously or temporarily, relative to control cultures, the difference lacked statistical significance (Table 6.2). As observed with D2763Kc2 cells, the median repeat change in D4132K was not significantly different between cultures continuously exposed to 5-azacytidine and replicates only treated over the initial 17 population doublings in culture ( $p=0.4712$ , two-tailed Mann Whitney  $U$  test).

Concluding, the exposure of *Dmt-D* cells carrying rapidly expanding CAG•CTG repeats to 5-azacytidine caused a statistically significant decrease in the median repeat gain. On the other hand, the effects of 5-azacytidine treatments on slowly contracting CAG•CTG repeats carried by D4132K cells were not as clear, and lacked statistical significance.

### **6.2.6. The effects of aspirin exposure on the dynamics of expanded CAG•CTG repeats**

NSAIDs, such as aspirin and sulindac, have previously been demonstrated to suppress the mutator phenotype associated with hereditary nonpolyposis colorectal cancer. Treatment of cell lines with known mutations in the MMR pathway with aspirin and sulindac reduced the microsatellite repeat instability normally associated with such cells. Aspirin is known to reduce the proliferative capacity of cell lines, produce a build up of cells at the  $G_0/G_1$  boundary and induce apoptosis (Ruschoff *et al.*, 1998). Given the established link between microsatellite instability and NSAIDs, aspirin would appear to be an excellent candidate chemical agent that might affect the metabolism of expanded triplet sequences.

Replicate cultures, derived from the D2763Kc2 cell line, were exposed to 5.6  $\mu\text{M}$  (1  $\mu\text{g/ml}$ ) aspirin (Table 6.1) and repeat size variability was assessed by SP-PCR techniques (Figure

6.6.A). The analysis showed an apparent reduction in the repeat expansion rate in cells cultures treated with aspirin (Figure 6.6.A). The significance of this result was confirmed by accurate single molecule sizing (Figure 6.6.B) and statistical analysis (Table 6.1). When replicate cultures of D4132K kidney cells were subjected to the same treatment (Table 6.2), a greater median rate of contraction was detected in aspirin treated cells (Figure 6.7) relative to control cultures, however the difference was not statistically significant (Table 6.2).

In summary, despite not being totally conclusive, the results described above suggest a possible stabilising effect of aspirin on rapidly expanding trinucleotide repeat tracts carried by *Dmt-D* kidney cultured cells.

### 6.3. Discussion

The lack of correlation between cell division rates, as assessed by population doubling time, and the levels of trinucleotide repeat instability in *Dmt-D* cultured cells (Chapter 3), implies that pivotal factors, other than DNA replication and cell division, must affect the dynamics of expanded repetitive sequences. The threat of excessive genetic changes needs constant attention as DNA becomes damaged by inherent errors in processes such as DNA replication, as well as through genotoxic stress from reactive cellular metabolites and exogenous stimuli (*e.g.* ionising radiation, UV light, chemical mutagens). Mammalian cells cope with the required monitoring and maintenance of genomic integrity by means of a complex network of DNA repair pathways and cell cycle checkpoints (Bartek and Lukas, 2001). In the view that the DNA repair mechanisms, associated with cell cycle checkpoints, are essential for the safeguard of genome integrity and DNA stability, it is conceivable that disruption of cell cycle progression, particularly of checkpoint controls, may have dramatic effects on the dynamics of inherently unstable trinucleotide repeats.

In an attempt to correlate the DNA repair efficiency with the levels of trinucleotide instability exhibited by *Dmt-D* cultured cells, the survival response of two cell lines derived from the same mouse was determined following exposure to UV-C radiation. The results suggested a parallel between higher levels of repeat instability in culture and greater sensitivity to UV light. If the increased death rates in kidney cells result from the accumulation of DNA damage, due to inefficient DNA repair, the differences between the two cell lines studied are consistent with an association between competent repair pathways and trinucleotide repeat stability. Alternatively, the greater survival of *Dmt-D* lung cells following exposure to UV light, might be due to the dysfunction of the apoptotic cascade (Van Sloun *et al.*, 1999), perhaps caused by an alternative MMR repair mechanism (Buermeyer *et al.*, 1999). Since mutations in mouse MMR genes, such as *Msh2*, *Msh3* and *Pms2*, and have been associated with reduced levels of somatic mosaicism (Manley *et al.*, 1999b; van Den Broek *et al.*, 2002; Chapter 8), it is reasonable to speculate that the greater trinucleotide stability in *Dmt-D* lung cells may be correlated with MMR mutations accumulated by this cell line, which would simultaneously explain greater survival to UV radiation and hydrogen peroxide, due to impaired activation of the apoptotic signal (Section 5.2.1).

UV light exposure induces cell cycle delay and activates DNA repair, primarily the NER pathway (Eckardt-Schupp and Klaus, 1999), therefore affecting DNA metabolism. To analyse the possible link between UV exposure and trinucleotide repeat dynamics in closer detail, *Dmt-D* cells were exposed to UV-C radiation, and the repeat size variability assessed by SP-PCR sensitive techniques. UV exposure resulted in altered repeat profiles, particularly in *Dmt-D* kidney cells that had previously exhibited high mutation rates under normal growth conditions. In contrast, *Dmt-D* lung cells, carrying stable transgenic sequences, failed to show any change detectable by SP-PCR techniques. Amongst the *Dmt-D* kidney cell lines analysed, the mutation trend, induced by UV exposure, varied between experiments, possibly indicating that it may be determined by the UV fluence used and nature of the cell line studied. While low doses of UV light (5 and 10 J/m<sup>2</sup>) appeared to induce an overall reduction in the average repeat number, higher UV fluences (20 J/m<sup>2</sup>) generated a dramatic expansion of the transgenic tract in *Dmt-D* cells after a single exposure. The contradictory results may be explained by different cellular responses stimulated by exposure to UV light. On one hand, low doses of UV-C radiation may primarily stimulate DNA repair mechanisms, mainly the NER pathway, which might alter the repeat dynamics in culture, through the alternative processing of mutation intermediates. On the other hand, higher levels of radiation may trigger an apoptotic pathway, leading to cell death and to changes in the repeat numbers detected in cultured cells, due to the preferential survival of cell subpopulations, which happen to carry a particular repeat size, either shorter or longer than the average. In addition, upon irradiation with high doses of UV (50-100 J/m<sup>2</sup>), p53 activates transcription of the human mismatch repair gene *MSH2* (Scherer *et al.*, 2000), and mutations in *msh2* increase the UV sensitivity of NER deficient yeast strains (Bertrand *et al.*, 1998). Studies of photoproduct binding by the human MSH2-MSH6 heteroduplex have demonstrated specific binding to a variety of mismatched CPDs and 6-4PP, but not to matched photoproducts (Mu *et al.*, 1997; Wang *et al.*, 1999a). MMR might therefore mediate the repair of UV-induced lesions when error-prone DNA synthesis passes UV photoproducts, producing photoproduct/base mismatches, in which damaged nucleotides are superimposed on mismatches (Liu *et al.*, 2000). If mouse Msh2-Msh6 heteroduplexes were involved in the excision of incorrect bases, inserted in nascent strands opposite photoproducts, altered levels of repeat length variability would be expected following UV exposure, given the reported roles of *Msh2* and *Msh6* as modifiers of trinucleotide repeat somatic mosaicism (Manley *et al.*, 1999b; van Den Broek *et al.*, 2002).

Interestingly, even when the dynamics of the repetitive transgenic tract was affected by UV radiation, the effect was not progressive and revealed to be more pronounced after a single exposure than following subsequent UV treatments. If a hypothetical radiation-induced repair pathway were permanently activated in culture, and inherited through cell division (*e.g.* epigenetic modification), this could possibly explain the absence of a progressive effect in cultures successively exposed to UV-C radiation. There is indeed evidence that a radiation-induced signal, which causes tandem repeat mutation in the mouse germline, is itself transmissible from parent to offspring, leading to the appearance of radiation-induced mutants in the germline of the unexposed progeny of irradiated mice (Dubrova *et al.*, 2000).

Complex mechanisms of adaptation exhibited by cells treated with UV might have accounted for the peculiar mutation profile exhibited by D2763K kidney cells. When repair of UV-induced lesions fails, the damaged cell may eventually resume its cell cycle, a phenomenon termed adaptation (Weinert, 1998). The generation of sufficient UV dimers to block the expression of genes involved in cell cycle arrest, such as p21, may mediate this process (McKay *et al.*, 1998). Assuming that D2763K cells exposed to one dose of 20 J/m<sup>2</sup> had the ability to adapt to UV exposure, the increased levels of expansion-biased trinucleotide repeat instability detected might be caused by the opportunity given to damaged cells to survive and generate genetically altered progeny. However, upon a second exposure to UV light these cells may accumulate numerous DNA lesions, so many that they become incompatible with cell survival, and the cells eventually die and are eliminated from the culture. Only the cells that did not show the capacity to adapt in the first place, and instead repaired their DNA, would be able to survive consecutive UV exposures, explaining the re-establishment of levels of repeat size variability identical to those of the progenitor cells, in cultures continuously exposed to UV light. The unusual response of D2763K cells to UV exposure could only be accounted for by adaptation, if the cells present in the original culture exhibited a considerable degree of variability in terms of their response to UV-induced lesions.

In summary, UV-C light may directly affect the repeat metabolism, maybe through the induction of DNA lesions and consequent activation of DNA repair pathways, which may lead to alternative processing of mutation intermediates. Alternatively, apparent changes in the trinucleotide repeat profile might be mediated by the preferential death of cells carrying particular repeats lengths. It should however be mentioned that while providing some insight into the mechanism of trinucleotide repeat biology, the exposure of *Dmt-D* cells to UV-C cells has little biological significance, as UV-C is effectively absorbed by the atmosphere.

Since altered trinucleotide repeat profiles were detected in *Dmt-D* cultured cells following DNA damage and cell cycle arrest, induced by exposure to UV light, disturbance of the cell cycle checkpoints may affect the dynamics of expanded CAG•CTG trinucleotide repeats. Forced progression of cultured cells through the G<sub>2</sub>/M DNA damage checkpoint led to a significant increase in the rates of trinucleotide repeat expansion in both cell lines studied. Despite the dramatic increase in the population doubling time in the presence of caffeine, the appearance of longer alleles in the presence of caffeine is consistent with the accumulation of non-repaired DNA replication slippage errors during mitosis, due to checkpoint override. Caffeine primarily inhibits the G<sub>2</sub>/M checkpoint, whose general purpose is to allow cellular recovery from DNA damage and prevent cell division in the presence of DNA damage (Blasina *et al.*, 1999). Structures in the replication complex or unreplicated DNA may send biochemical signals to inhibit mitotic entry (Elledge, 1996). Since CAG•CTG repeats can adopt non-B-DNA conformation (Chastain *et al.*, 1995; Pearson and Sinden, 1996; Pearson and Sinden, 1998a; Pearson and Sinden, 1998b; Sinden, 1999) and inhibit DNA polymerisation *in vitro* (Kang *et al.*, 1995b), it is reasonable to assume that, if these unorthodox DNA structures also exist *in vivo*, they may trigger a biochemical signal that arrests cell cycle progression, allowing the cell to repair its DNA. Treatment with caffeine may

inhibit such a checkpoint and induce the accumulation of mutations within the trinucleotide repetitive tract, which may presumably result from the inability to correct and convert loop-outs and hairpins formed within the transgene, into more conventional B-DNA structures.

Despite being forced to enter mitosis, cell populations exposed to caffeine still grow slower than controls, given the inhibitory effect of caffeine on DNA replication (Deplanque *et al.*, 2001; Deplanque *et al.*, 2000) or the high apoptotic levels induced by the continuous accumulation of replication errors and double strand breaks (Norbury and Hickson, 2001). Therefore, the effects of caffeine on the dynamics of CAG•CTG repeats may be alternatively explained by the decreased rate of DNA replication, and slowing down of the S phase of the cell cycle.

However, when DNA replication was decelerated by exposure to araC, which does not induce checkpoint override, *Dmt-D* kidney cells exhibited lower levels of repeat expansion than those observed in control cultures, in contrast with the results described for caffeine treated cells. This result suggests that the changes in CAG•CTG repeat dynamics detected in cells exposed to caffeine cannot be explained by a slow cell progression through the S phase of the cell cycle. Instead, the increase in the rates of expansion observed in cells treated with caffeine may be intimately connected with the override of the G<sub>2</sub>/M DNA damage checkpoint.

The uncoupling of the DNA synthesis on the two strands, induced by araC treatment, generates asymmetric replication forks, with areas of single stranded DNA corresponding to the site of blockage of DNA synthesis on the leading strand. This phenomenon may have biologically relevant consequences. The imbalanced DNA synthesis induced by araC is often associated with the generation of genomic instability (Carbone *et al.*, 2001).

The cytidine analogue may affect the repeat metabolism by means other than the slowing down of DNA replication *per se*. T-lymphoblastoid cell lines incorporate araC into their DNA, not only by replication DNA synthesis, but also by repair synthesis. In fact, this cytosine analogue is incorporated into repairing DNA to a greater extent than into the replicating DNA of proliferating cells (Iwasaki *et al.*, 1997). Indeed, araC is known to sensitise murine and human cells to a variety of mutagens, including radiation and radiomimetic compounds. This sensitisation is due to the ability of araC to act as a cytosine analogue and to inhibit DNA synthesis during DNA repair (Fram and Kufe, 1985). It is therefore reasonable to assume that the inhibition of DNA repair may, at least in part, account for the changes in the expansion rates observed in the presence of araC, again suggesting that the repair mechanisms have a dramatic impact on the mutation mechanism of expanded trinucleotide repeats.

While araC is incorporated into DNA to a limited extent, because of its inhibitory effect on DNA synthesis, 5-azacytidine is extensively incorporated into newly synthesised DNA strands. Previous correlations were reported between trinucleotide repeat instability, GC content and proximity to CpG islands (Brock *et al.*, 1999). A potential way in which this effect could be mediated is by DNA methylation at CpG dinucleotides. When exposed to 5-azacytidine, at concentrations that do not affect cell proliferation as assessed by the population doubling time, D2763Kc2 cultures exhibited a reduction in the median rates of trinucleotide repeat expansion, relative to the control cells. The results presented here indicate that DNA methylation may act as a

modifier of trinucleotide repeat instability *in vivo* and raise the possibility that it may also account for interlocus differences and tissue-specific patterns of somatic mosaicism. Given the major impact of DNA methylation on DNA metabolism, there are several mechanisms by which hypomethylation, induced by 5-azacytidine, may stabilise the trinucleotide repeat tract. Methylation plays a pivotal role in establishing and maintaining the inactive state of a gene by rendering the chromatin structure inaccessible to the transcription machinery (Bird and Wolffe, 1999; Razin, 1998). Therefore, DNA hypomethylation induced by 5-azacytidine leads to the expression of previously silent genes (Bird and Wolffe, 1999; Momparler and Bovenzi, 2000). Methylation of the promoter region of the mismatch repair gene *MLH1* correlates with its lack of expression in primary colon tumours, whereas normal adjacent tissue and colon tumours that expressed this gene did not show signs of *MLH1* promoter methylation (Veigl *et al.*, 1998; Wheeler *et al.*, 1999a). Expression of *MLH1* in colorectal cancer cells was restored after treatment with 5-aza-2'-deoxycytidine (Deng *et al.*, 1999; Herman *et al.*, 1998). Similarly, both *MHS6* mRNA and protein levels were increased by 5-azacytidine treatment in human cell lines that did not express detectable levels of *MSH6* (Bearzatto *et al.*, 2000). Therefore, *MSH6* joins *MLH1* MMR gene as a potential candidate gene whose expression can be reactivated by hypomethylation (Bearzatto *et al.*, 2000). The possibility of enhanced expression of *Msh6* in D2763Kc2 cells upon 5-azacytidine exposure is of particular interest, since *Msh6* acts as an inhibitor of trinucleotide repeat somatic mosaicism (van Den Broek *et al.*, 2002), and may therefore be correlated with the stabilisation of the transgenic tract observed in these cells. 5-Azacytidine may affect the efficiency of the MMR pathway, not only by reactivating dormant genes, but also by interfering with the mechanism of strand discrimination. In *E. coli* DNA MMR is directed to the newly synthesised strand due to its transient lack of adenine methylation (Modrich, 1997). Although the nature of the strand signal has not been identified in any eukaryotic organism, heteroduplex DNA containing a site-specific incision is subject to mismatch-provoked, strand-specific repair by human cell extracts. Following DNA replication, mammalian DNA is characterised by a transient, strand-specific CpG hemimethylated parental strand. Cytosine hemimethylation within CpG sites may represent an analogous mechanistically plausible means of targeting mismatch correction (Cleaver, 1994). In an effort to identify human proteins that may mediate strand discrimination during MMR, methyl-CpG binding endonuclease 1 (MED1) was identified through its interaction with MLH1 protein in a yeast two-hybrid system analysis (Bellacosa *et al.*, 1999). MED1 binds with different affinity to methylated and hemimethylated CpG sequences, hence discriminating between the two methylation status. Furthermore, deletion of the methyl-CpG binding domain of MED1 is associated with microsatellite instability (Bellacosa *et al.*, 1999). MED1, also known as MBD4 (methyl-CpG binding domain 4), has also been identified in mice, binding specifically to methyl-CpG *in vitro* and *in vivo* (Hendrich and Bird, 1998). Similar to the bacterial methyl-directed reaction, strand specificity in human MMR could be mediated, at least in part, by MED1-dependent recognition of transiently hemimethylated CpG sites generated after DNA replication (Bellacosa *et al.*, 1999). However, neither hemimethylation nor full methylation at CpG sequences was found to have an effect on the magnitude or rate of correction of mismatched simian virus 40 (SV40) DNA

molecules by nuclear extracts, prepared from HeLa cells, questioning a possible role of DNA methylation in eukaryotic strand discrimination (Drummond and Bellacosa, 2001). Nevertheless DNA methylation has been reported to play an important role in maintaining genomic integrity (Chen *et al.*, 1998b). If we assume that DNA methylation has some influence on MMR efficiency, the reduced repeat instability detected in *Dmt-D* kidney cells upon exposure to 5-azacytidine might be mediated by the lack of repair activity on mismatches, possibly generated by replication slippage (Manley *et al.*, 1999b; van Den Broek *et al.*, 2002; Chapter 8).

Nevertheless, other factors must also be considered. By modifying the DNA methylation status, the incorporation of 5-azacytidine into genomic DNA can also change the chromatin structure and the physical properties of DNA, as well as the protein/DNA interactions (Hendrich *et al.*, 1999; Hendrich and Bird, 1998; Takahashi-Hyodo *et al.*, 1999). All these modifications may account for the effects of 5-azacytidine on the dynamics of triplet repeats. However, if the decreased levels of repeat instability in *Dmt-D* cultures cells treated with 5-azacytidine were the direct consequence of the incorporation of this antimetabolite, rather than the outcome of hypomethylation, a greater effect would be expected in cells continuously exposed to the chemical, compared to cells that were only temporary treated. That was not the case, as both treatments, resulted in similar rates of expansion, both significantly lower than in the controls. Hypomethylation *per se* can also affect chromatin structure and protein accessibility (Kass *et al.*, 1997; Razin, 1998). DNA structural changes, and altered DNA/protein interactions induced by global genomic hypomethylation may therefore account for the decreased rate of expansion observed in treated cells.

In addition, both araC and 5-azacytidine have minor effects on topoisomerase activity. Once incorporated into the DNA, 5-azacytidine may change the topoisomerase II cleavage sites, altering the binding and activity of this enzyme (Jablonka *et al.*, 1985; Jackson-Grusby *et al.*, 1997; Jones and Taylor, 1980; Lopez-Baena *et al.*, 1998; Santi *et al.*, 1983; Takahashi-Hyodo *et al.*, 1999). AraC traps the topoisomerase I cleavage complexes, preventing the DNA religation step (Pourquier *et al.*, 2000). Since topoisomerases appears to be key mediators of trinucleotide repeat somatic mosaicism (Chapter 4), the stabilising effect of both cytidine analogues on expanded triplet repeat sequences may be explained by a reduced activity of topoisomerase activity. Similarly, the outcome of UV exposure on the dynamics of unstable trinucleotide repeats may be intimately associated with modifications of DNA topology. A subset of proteins likely to be involved in the repair of UV photolesions are redistributed along the chromatin upon UV irradiation, being translocated from a loose to tight association with genomic DNA (Otrin *et al.*, 1997). This may influence the DNA topology and therefore the dynamics of expanded CAG•CTG repeats in *Dmt-D* cultured cells (Chapter 4).

In addition to all the mechanisms by which the genotoxic discussed above may modify the dynamics of trinucleotide sequences, apoptosis may also play a role, determining which cells survive and pass their genetic information onto the progeny. NSAIDs, including aspirin, inhibit the growth of tumour cells, through the induction of apoptosis and inhibition of DNA replication. Nevertheless, inhibition of DNA replication must play a minor role, as DNA synthesis in

carcinoma cell lines and Swiss 3T3 mouse fibroblasts never dropped below 50% of that in controls, when treated with doses similar to those used in this study (Castano *et al.*, 1997; Ricchi *et al.*, 1997; Richter *et al.*, 2001). D2763Kc2 kidney cells chronically exposed to aspirin, at concentrations that did not inhibit cell proliferation, as assessed by population doubling times, showed a significant reduction in the expansion rates of expanded CAG•CTG trinucleotide sequences. However, the exposure of D4132K cultures, failed to reveal a significant effect on the dynamics of the transgenic repeats. Interestingly, aspirin also stabilises microsatellite sequences in tumour cell lines derived from HNPCC patients carrying mutations in one of the following MMR genes: *MLH1*, *MSH2* and *MLH6* (Ruschoff *et al.*, 1998). Despite the reduction in the proliferative rates in aspirin treated cells and the altered cell cycle distribution, the reduction in dinucleotide frequency of mutation cannot be a mere consequence of reduced proliferation rate, as no significant changes in the frequency of microsatellite instability was detected when slowly growing serum-starved cells were compared with cells growing under standard conditions (Ruschoff *et al.*, 1998). Moreover, the amount of aspirin that *Dmt-D* cells were exposed to, although sufficient to stabilise expanded trinucleotide repeat tracts, did not affect the population proliferation rates, as assessed by population doubling times. Rather than being mediated by a slow growth rate, the stabilisation of dinucleotide and trinucleotide microsatellites by aspirin might be due to the induction of apoptosis, which mediates a genetic selection for microsatellite stability (Ruschoff *et al.*, 1998). The mechanism underlying such selection is currently unknown, and levels of apoptosis should be measured in *Dmt-D* cells in order to confirm the basis for the trinucleotide stabilisation by aspirin in these cultures. If increased apoptotic levels were confirmed in *Dmt-D* cells exposed to aspirin, one could speculate that aspirin might alter the recognition of the threshold of genetic instability, as well as the subsequent apoptotic decision, and trigger apoptosis in cells that exhibit higher levels of trinucleotide repeat instability, thereby selecting for cells that carry stable repeat tracts.

It has been recently reported that aspirin inhibits hydrogen peroxide-induced DNA strand breaks in plasmid DNA (Hsu and Li, 2002). Possibly, the inhibition of oxidative DNA damage by aspirin in *Dmt-D* mouse cells may also contribute to the stabilisation of trinucleotide repeat tracts in culture.

The expression of a large number of genes (more than 170), associated with transcription, signal transduction, cell cycle regulation and apoptosis is influenced by aspirin in colon cancer cell lines (Iizaka *et al.*, 2002), making it very difficult to pinpoint the actual cause of the stabilisation of expanded CAG•CTG repeats in cultured *Dmt-D* kidney cells. MMR gene expression (Iizaka *et al.*, 2002) and protein levels (Ruschoff *et al.*, 1998) could not account for the microsatellite stabilisation observed, eliminating expression compensation as the cause of the effect observed. Nevertheless, altered expression of unidentified genes was reported upon aspirin treatment (Iizaka *et al.*, 2002), and the possibility exists that these may either indirectly affect or be involved in DNA repair.

It is worth noting that the lack of statistical significance in the results presented for the 5-azacytidine and aspirin treatments on D4132K cultures may be explained by the intrinsic behaviour of the repeats in these cells. First, the repeats carried by this cell line do not show a tendency for

rapid expansion, as the repeats in the D2763Kc2 cell line. Second, the transgenic alleles carried by D4132K cells fall into two distinct peaks of repeat length variability. Therefore, only an extensive single molecule analysis, performed for each one of the two cell subpopulation observed, could resolve the problem and clarify the effect of 5-azacytidine and aspirin, if any, on the dynamics of the expanded CAG•CTG repeats carried by the D4132K cell line.

In summary, drug treatments that resulted in checkpoint override resulted in destabilisation of trinucleotide repeats, whereas, treatments that slowed down cell proliferation by inhibition of DNA replication, were associated with the stabilisation of the repetitive tracts. Global genomic hypomethylation was also found to stabilise trinucleotide repetitive tracts, as well as exposure to aspirin. Taken together, the results presented here provide strong evidence for the involvement a complex network of factors, acting either in *trans* or *cis*, in the molecular mechanism of trinucleotide repeat mutation. Whatever the precise mechanism of action, some environmental agents destabilise trinucleotide repeat tracts and induce enhanced levels of repeat instability. If a similar scenario occurred in humans, environmental agents might possibly accelerate the development and progression of the symptoms. Caffeine itself is unlikely to be considered amongst these harmful agents, due to the high (millimolar) concentrations required to elicit checkpoint inhibitory activity. On the contrary, the ability of certain chemicals to specifically reduce the expansion rates in expanded CAG•CTG trinucleotide repeats carried by cultured *Dmt-D* cells, presents the possibility that clinically relevant drugs may provide major therapeutical benefits, by stabilising the repeats at loci associated with severe human diseases.

## 7. Trinucleotide repeat dynamics in slowly proliferating and non-dividing *Dmt-D* cells

### 7.1. Introduction

The first model proposed to address the molecular mechanisms driving repeat instability, predicted that small expansions or deletions within tracts of trinucleotide repeats were generated during genome replication, by DNA polymerase slippage, as cells divide. During DNA replication, a double-stranded DNA helix must first be unwound and unzipped; the two separated strands are then used as templates for the production of the two new strands. When the DNA includes long tracts of simple repeats, there is the potential for the growing DNA chain and the original template to become misaligned if DNA replicating enzymes pause and separate from the template. This can lead to deletions or expansions of sequences when replication reinitiates (Richards and Sutherland, 1994). Despite its wide acceptance, this hypothesis raises several questions, which have not yet been resolved.

First, if trinucleotide repeat instability were primarily the result of replication slippage, mutations in MMR genes should increase the mutation rate. However, the study of mouse models carrying unstable transgenic CAG•CTG sequences, has revealed lower levels of somatic mosaicism in mice deficient for *Msh2* (Manley *et al.*, 1999b), *Msh3* (van Den Broek *et al.*, 2002) and *Pms2* (Chapter 8) repair genes.

Second, if polymerase slippage is the driving force of repeat somatic instability, why is it that tissues with higher cell turnover rates, such as blood, do not necessarily show higher degrees of trinucleotide repeat mosaicism? In DM1 patients, allele sizes in post-mitotic muscle cells are consistently larger than in rapid proliferating circulating lymphocytes (Anvret *et al.*, 1993; Ashizawa *et al.*, 1993; Lavedan *et al.*, 1993). Polyacrylamide gel electrophoresis of bulk DNA PCR amplification products, has initially suggested the absence of extreme mosaicism in human HD tissues such as blood, liver and intestine, where cell turnover is high, indicating that cell division alone is not the only event promoting somatic instability (Telenius *et al.*, 1994). In addition, in SBMA patients the most prominent trinucleotide repeat somatic mosaicism is observed in the cardiac and skeletal muscles, composed predominantly of post-mitotic cells; and the skin, one of the tissues with the highest cell turnover rate, does not show a detectable increase in somatic repeat instability with age (Tanaka *et al.*, 1999), strongly suggesting that cell division is unlikely to be the major cause of repeat size mutation.

In mouse models that recreate CAG•CTG somatic mosaicism, repeat instability occurs at different rates in distinct tissues, depending on the chromosomal context (Fortune *et al.*, 2000; Lia *et al.*, 1998; Mangiarini *et al.*, 1997; Seznec *et al.*, 2000), and showing continued repeat expansion as the mice age (Fortune *et al.*, 2000; Kennedy and Shelbourne, 2000; Lia *et al.*, 1998; Mangiarini *et al.*, 1997; Seznec *et al.*, 2000). However, in agreement with previous observations in humans, it

appears that repeat variability can occur in the absence of the DNA replication that is associated with genome duplication. Indeed repeat size mutations often occur to a greater extent in cells that do not divide (such as brain and kidney cells) (Fortune *et al.*, 2000; Kennedy and Shelbourne, 2000; Lia *et al.*, 1998; Seznec *et al.*, 2000). In both *Dmt-D* mice (Monckton *et al.*, 1997) and in a knock-in HD mouse model (Shelbourne *et al.*, 1999) high levels of somatic mosaicism were detected in the striatum by sensitive SP-PCR techniques (Fortune *et al.*, 2000; Kennedy and Shelbourne, 2000). Given that the vast majority of cells in the striatum are neurons, and therefore post-mitotic, it is difficult to reconcile this observation with small, successive replication-based expansions. In fact, the mutation profile obtained with a sample of striatal tissue containing replication-competent cells derived from the subventricular zone, did not differ significantly from that obtained when the same cells were excluded from the analysis (Kennedy and Shelbourne, 2000). CAG•CTG repeat length mutation has also been reported to occur in mature oocytes from a SCA1 transgenic mouse model (Kaytor *et al.*, 1997) and in the developing sperm of R6/1 mice (Kovtun and McMurray, 2001), at stages when DNA replication and recombination do not take place. Both replication and recombination are thereby excluded as obligatory pathways to repeat variability in germ cells.

However, it should be considered that most of these studies have been predominantly performed on whole human or mouse tissue samples, consisting of multiple cell types likely to have differing dynamics in terms of both cell turnover rates and repeat instability. The presence of distinct cell populations most likely complicated attempts to establish simple correlations and may have masked some subtle relationships. When differences in repeat expansion among neuronal subgroups in DRPLA brains were revealed by single-cell PCR techniques, the distinct levels of instability detected could not be simply accounted for by the number of cell divisions (Hashida *et al.*, 2001). In addition, monitoring the repeat dynamics in homogenous *Dmt-D* cell lines over long periods of time, failed to reveal any correlation between rates of cell turnover, assessed by population doubling times, and repeat instability (Chapter 3). In particular, the extremely low mutation rates observed in a very rapidly proliferating *Dmt-D* lung cell line, may reflect independence of repeat stability on factors other than DNA replication and/or cell division.

In summary, circumstantial evidence gathered from the study of somatic mosaicism, either in humans or mouse models, revealed that specific tissues have a non-random level of CAG•CTG instability, apparently not related to the number of cell divisions alone, strongly suggesting that there are separate factors that influence somatic triplet instability. It is reasonable to speculate that at least some mutants may arise during the resting phase of the cell cycle, in a time-dependent fashion. In order to test this hypothesis directly, it would be ideal to observe the accumulation of repeat length variation in non-dividing cells over time. Several strategies exist to arrest cell division in culture, such as contact inhibition, serum starvation, and chemical treatment with genotoxic agents that block cell cycle progression.

High cell density induces cell cycle arrest. This phenomenon, termed contact inhibition, is generally observed in normal or untransformed cells. Most non-transformed cell lines respond to confluence by arresting the cell cycle in a viable G<sub>1</sub> phase, whereas immortalised cell lines growing

in monolayer do not stop mitotic progression in response to high cell density and are subjected to density-dependent apoptosis (Brezden and Rauth, 1996).

Cell cycle progression is controlled by the periodic activation of cyclin-dependent kinases (CDKs). CDKs become activated by their association with activating subunits, referred to as cyclins. CDK activity is also modulated through association with negative regulatory subunits, known as CDK-inhibitory proteins, which prevent cells from progressing through the cell cycle. These negative regulators include a variety of proteins such as p15, p16, p19, p21, p27 and p57 (Elledge, 1996). In response to contact inhibition, both mRNA and protein levels of p15, p16 and p27 are up-regulated as the cell density increases; whereas cyclins A and E are down-regulated, leading to inactivation of CDKs and cell cycle arrest (Chen *et al.*, 2000; Kato *et al.*, 1997; Yanagisawa *et al.*, 1999). Sustained accumulation of p21 in the nucleus was also reported in contact-inhibited cultures, suggesting that targeting and maintaining p21 in the nuclear compartment may contribute to slowing or arresting cell cycle progression, once cell-to-cell contacts are established. In addition, p21 may also play a role in inducing apoptosis (Ritt *et al.*, 2000). The additive effects of contact inhibition and serum starvation on [<sup>3</sup>H]thymidine incorporation by human fibroblasts led to the conclusion that the mechanisms underlying cell cycle arrest might be different (Dietrich *et al.*, 1997). A strong decrease in CDK2 activity has also been reported in serum-starved cells, while the protein levels of p27 and p16 remain low. In contrast, a rapid decrease of cyclin D1 and cyclin D3 was observed, which does not occur in contact-inhibited cells (Dietrich *et al.*, 1997).

Mitomycin C is an antibiotic isolated from *Streptomyces*, which acts as a bifunctional DNA alkylating agent, inducing the formation of covalent links between guanine residues in complementary DNA strands (Iyer and Szybalski, 1963; Waring, 1968). The polycyclic structures might favour intercalation of this compound into DNA, as a step prior to selective crosslinking (Iyer and Szybalski, 1963). The remainder of the bound mitomycin C, which does not participate in the formation of interstrand crosslinks, presumably reacts by monofunctional alkylation (Waring, 1968). Both crosslinks and mono-adducts are important cytotoxic lesions. The bis-adduct fit tightly into the minor groove of DNA and produces minimal distortion of the B-DNA helical structure (Tomasz *et al.*, 1987), being therefore difficult to recognise and repair, and very toxic. The presence of a crosslink between the parental DNA strands effectively prevents the progress of the replicating fork, and thus might account for the powerful blockage of cell division (Iyer and Szybalski, 1963; Waring, 1968). Consequently, mitomycin C is a potent antiproliferative agent against mammalian cells, preventing DNA replication until the alkylated residue is replaced by intracellular DNA repair mechanisms (Tomasz *et al.*, 1987). Cell cycle analysis has revealed an accumulation of cells arrested in S and G<sub>2</sub>/M phase following mitomycin C treatment, through the up-regulation of p53 and p21 (Kang *et al.*, 2001). Mitomycin C has the distinct theoretical advantage that administration of a single dose may inhibit mammalian DNA replication and cell proliferation for at least three to four weeks (as long as measured in previous studies), probably suggesting permanent DNA damage without recovery (Castaneda and Kinne, 1999; Ho *et al.*, 1997; Jampel, 1992; Murayama *et al.*, 1996; Takahashi *et al.*, 1998). Furthermore, the cytotoxicity of

mitomycin C is independent of the proliferative status of the cultures, with similar toxicities being measured in exponentially growing and plateau phase mouse mammary tumour cells (Rockwell and Hughes, 1994).

Apicidin, is a potent inhibitor of histone deacetylase that exhibits a broad spectrum of antiproliferative activity towards various cancer cell lines (Han *et al.*, 2000). The antiproliferative activity of apicidin in human cells is accompanied by morphological changes, cell cycle arrest at G<sub>1</sub> phase, and accumulation of histone hyperacetylation *in vivo*, as a result of the inhibition of histone deacetylase activity (Han *et al.*, 2000; Kim *et al.*, 2001). Several genes contain regulatory elements sensitive to acetylation, and are directly transcriptionally up-regulated in cells treated with histone deacetylase inhibitors. It is believed that transcription factors capable of activating transcription of acetylation-sensitive genes are present in the nucleus, but have a reduced affinity for their DNA binding sites when the local chromatin is deacetylated. Treatment with inhibitors of histone deacetylase activity shifts the equilibrium in the direction of acetylation, resulting in binding of transcriptional activators and subsequent gene transcription. Virtually all genes regulated by acetylation levels are associated with cell growth, differentiation or development (Hassig *et al.*, 1997). Whereas the protein levels of cyclin D1, CDK2 and p53 are not affected by apicidin exposure, the expression of p21 is markedly up-regulated by apicidin (Han *et al.*, 2000; Kim *et al.*, 2001). Activation of p21 transcription has been associated with growth arrest and induction of differentiation in response to various agents (Dotto, 2000; Gartel and Tyner, 1999), and may therefore be the main cause of the antimitogenic properties of apicidin.

Interferons constitute a family of secreted polypeptides, also known as cytokines, of which interferon  $\alpha$  is an example. Interferons are capable of modulating a variety of cellular responses, including proliferation of both malignant and non-malignant mammalian cells of many different origins (Sangfelt *et al.*, 2000). However, cells differ greatly in their sensitivity to the effects of interferons, with some cells being extremely sensitive, and others being more or less resistant. Treatment of susceptible cells with interferon  $\alpha$  has been found to affect various phases of the mitotic cycle. Most commonly, exposure to interferon  $\alpha$  leads to G<sub>1</sub> arrest, although sometimes slowed growth is due to a blockage or prolongation of S-phase or a lengthening of all cell cycle phases (Sangfelt *et al.*, 2000). Interferon  $\alpha$  can exert profound anti-mitotic effects in cell cultures, most likely involving multiple molecular pathways by regulating several components directly involved in the cell cycle machinery (Sangfelt *et al.*, 2000). This cytokine has been shown to strongly repress the activity of CDKs, reduce the expression levels of cyclins A and D3 and to induce the activity of p15, p19, p21 and p27 in a number of different cell types, correlating with cell growth inhibition (Matsuoka *et al.*, 1998; Sangfelt *et al.*, 1999; Sangfelt *et al.*, 1997; Sangfelt *et al.*, 2000; Tiefenbrun *et al.*, 1996).

*Dmt-D* kidney cells, exhibiting high levels of repeat instability, were cultured in low serum or contact-inhibition conditions, and exposed to mitomycin C, apicidin or interferon  $\alpha$ , in an attempt to stop culture proliferation. The effects of cell cycle arrest or high cell density conditions on somatic mosaicism were assessed by sensitive techniques, in order to gain some insight on the

dependence of the mechanism of trinucleotide repeat length mutation on cell division and DNA replication.

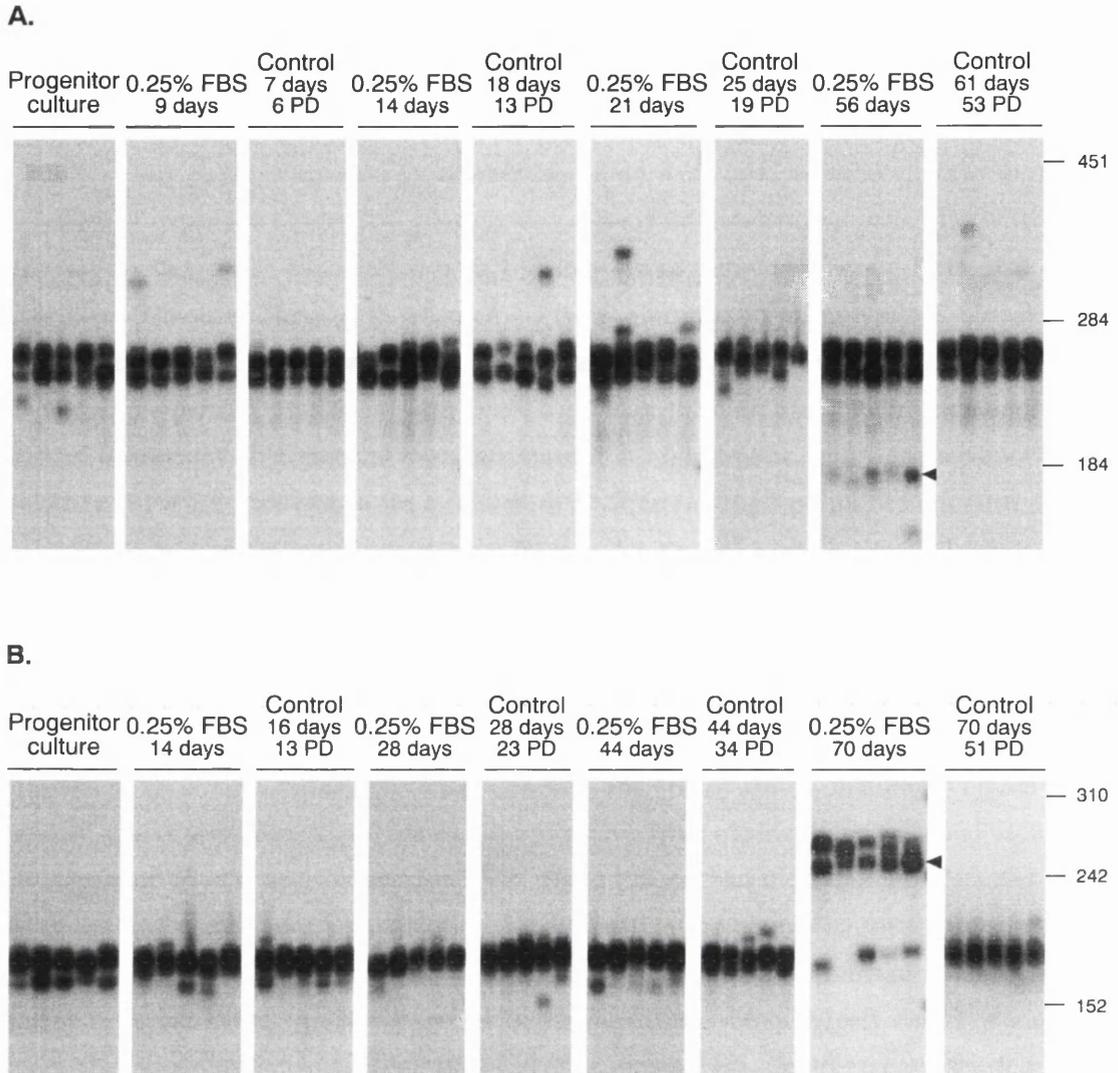
## 7.2. Results

### 7.2.1. *Trinucleotide repeat dynamics in Dmt-D cells growing at high density in low foetal bovine serum*

In order to check for the possibility that non-dividing cells can still accumulate repeat length mutations, *Dmt-D* kidney cells were arrested by serum starvation. Four D2763K and D3111K replicate cell cultures (Table 3.2) were maintained in standard growth medium, supplemented with 10% (v/v) foetal bovine serum (FBS), until they reached 80-90% confluency and the FBS levels were then decreased to 0.25% (v/v) to inhibit cell proliferation. Nevertheless, the accumulation of multiple cell layers over time was observed by phase contrast microscopy, suggesting continuing cell proliferation, still possible even under conditions of low FBS. In addition, the culture medium became acidic very quickly, as revealed by the development of a yellow/orange colour, indicating a rapid cell metabolism in serum-starved cultures.

Cells maintained at high density in 0.25% (v/v) FBS, without subculturing, were harvested at different time points and the levels of repeat size variability assessed by high DNA input SP-PCR techniques, and compared to those detected in the progenitor cells and in proliferating control cultures, maintained in 10% (v/v) FBS for similar periods of time (Figure 7.1). None of the control cultures exhibited major trinucleotide repeat instability, with no major expansion or deletion mutants being detected. The repeat dynamics in serum-starved D2763K did not reveal major differences relative to the controls. However, a small subpopulation of shorter alleles, showing clear evidence of clonal expansion, was detected at 56 days. In contrast, equally short repeats were not detected in the corresponding control culture (Figure 7.1.A). Similarly, D3111K confluent cultures, maintained in low FBS, did not exhibit major differences in the degree of repeat length variation relative to the controls, up to 70 days in culture. By this time point, a population of cells carrying longer repeats, and again showing clear signs of clonal proliferation, overgrew the culture (Figure 7.1.B).

The accumulation of subpopulations of mutant alleles, showing little repeat length variation, in *Dmt-D* cells cultures grown in 0.25% (v/v) FBS, is indeed consistent with the clonal expansion of a few mutant cells exhibiting enhanced survival to conditions of high density, and rejects the hypothesis that these cultures were truly arrested by serum starvation. Nevertheless, it is not unreasonable to assume that serum-starved cells were indeed proliferating at lower rates than control cultures, as they were not passaged for at least 56 days.



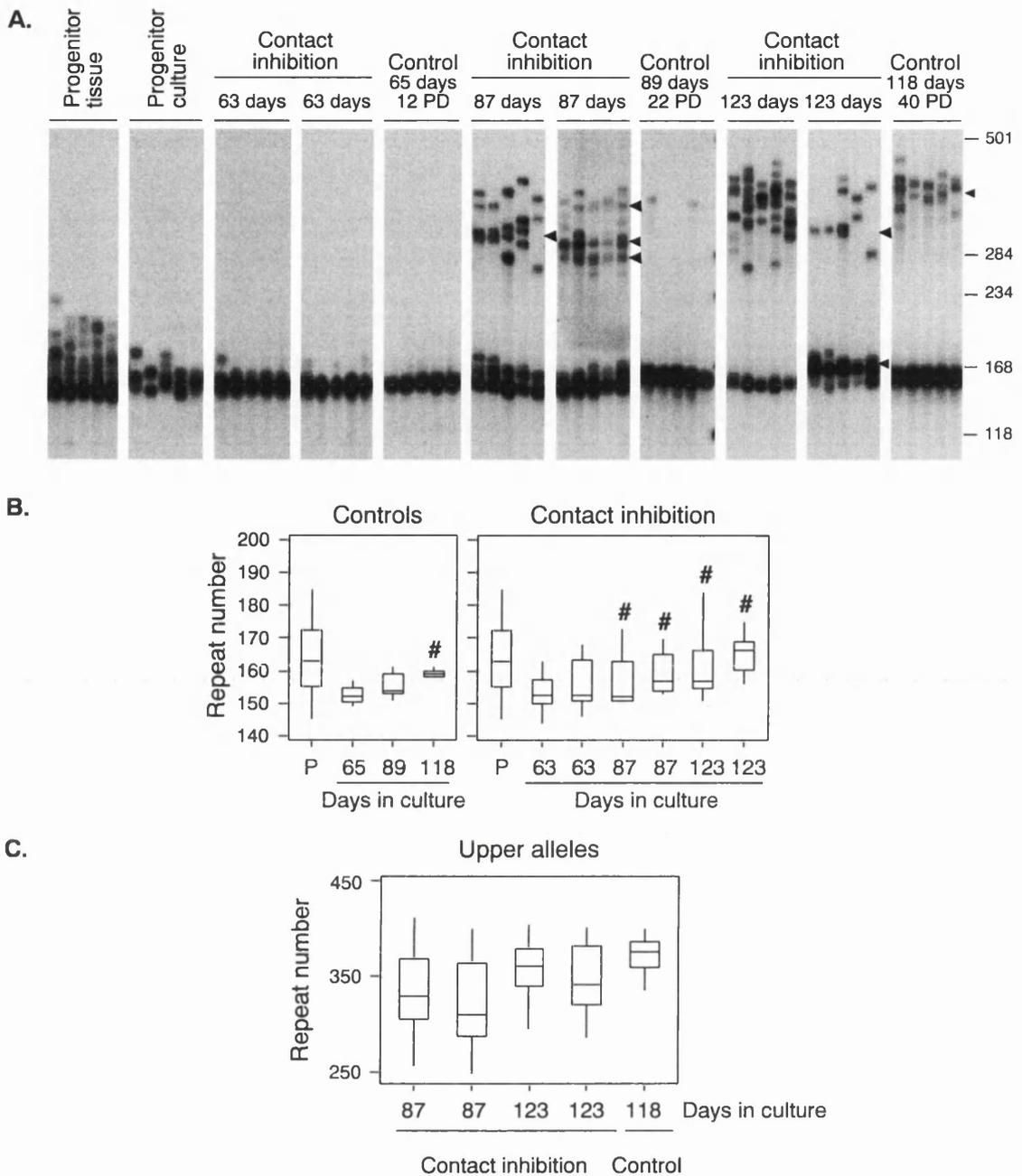
**Figure 7.1. Trinucleotide repeat dynamics in *Dmt-D* kidney cells maintained at high density by serum starvation.**

The autoradiographs show representative SP-PCR amplifications of DNA samples extracted from highly confluent *Dmt-D* kidney cell cultures, maintained in 0.25% (v/v) FBS. Fresh medium supplemented with 0.25% (v/v) FBS was added to the confluent cultures every two to three days. Control cells were cultured under standard growth conditions for the same period of time, and passaged when confluent. The number of population doublings (PD) for each control sample is shown in the figure. Two kidney cell lines were studied: **(A)** D2763K (derived from a six-month-old *Dmt-D* mouse), and **(B)** D3111K (established from a five-week-old *Dmt-D* mouse). Ten to 30 transgene molecules were amplified in each reaction. The molecular weight markers, converted into CTG repeat numbers are shown on the right. Black arrowheads (◄) indicate evidence of clonal expansion.

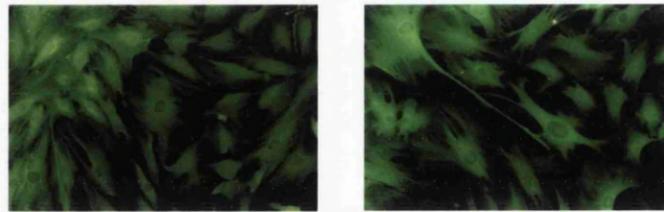
### **7.2.2. Trinucleotide repeat dynamics in confluent Dmt-D cell cultures arrested by contact inhibition**

Since cultured mouse cells undergo spontaneous immortalisation (Meek *et al.*, 1977), primary kidney cultures arrested by contact inhibition were used as an alternative to monitor the dynamics of expanded CAG•CTG trinucleotide repeats in non-dividing cells. D4393K cells were collected by enzymatic dissociation from a kidney tissue sample harvested from a six-month-old male *Dmt-D* mouse. These cells showed a typical fibroblastic phenotype, as revealed by their spindle morphology, vimentin expression and lack of staining for a panel of cytokeratins (Figure 7.3). Multiple parallel replicates were established after the first passage and expanded until they reached confluency. The cultures were maintained at high density, and fresh medium was added regularly. SP-PCR analyses were performed at different time points, to monitor repeat size variability in two arrested replicate cultures (Figure 7.2.A). The analysis revealed minor differences over the first 63-65 days in culture, but a dramatic accumulation of longer alleles was detected in arrested cells at later stages. A similar subpopulation of cells carrying longer alleles was not detected in control cells, growing for 89 days and 22 population doublings, but a few expanded mutants were observed at high DNA concentrations. However, similar expanded alleles were later observed in the control culture grown for 118 days and 40 population doublings (Figure 7.2.A). To quantify this observation, 20-50 transgene molecules from each culture were individually sized. The expanded alleles were on average ~300 repeat longer than those initially detected in culture. Given the low degree of repeat length variability within the population of longer alleles (Figure 7.2.A, black arrowheads), they were considered to have resulted from the clonal expansion of mutant cells exhibiting greater survival under conditions of high cell density, and analysed separately. Single molecule analysis of the shorter repeat tracts ( $\leq 240$  repeats) carried by control cells revealed the progressive expansion of trinucleotide repeat tracts from day 65 to day 118 (median gain of 0.132 repeats per day). Arrested cells also showed some degree of repeat expansion, but the increase in repeat number was not always statistically significant (Figure 7.2.B). Single molecule analysis of the longer alleles ( $> 240$  repeats) carried by arrested cells (Figure 7.2.C) revealed a dramatic and statistically significant increase in the median repeat number from day 87 to day 123 ( $p=0.0001$ , two-tailed Mann-Whitney *U* test).

Despite evidence suggesting that the longer repeats resulted from the clonal expansion mutant cells, rather than the accumulation of repeat size variation in non-dividing cells, extensive development of multiple cell layers was not observed using phase contrast microscopy. Therefore, a 5'-bromo-2-deoxyuridine (BrdU) incorporation and detection analysis was performed on D4393K cells hypothetically arrested by contact inhibition. Two incubation periods with BrdU were used. A 15-minute BrdU incubation period should only label the proportion of cells that are actively undergoing DNA replication during S phase of the cell cycle (also known as the "labelling index"). On the other hand, a 30-hour incubation should label most, if not all, of the cells that are still capable of active proliferation (also known as the "growth fraction") (Rew and Wilson, 2000). Despite the low plating efficiency of arrested cells onto an eight-well chamber slide, following

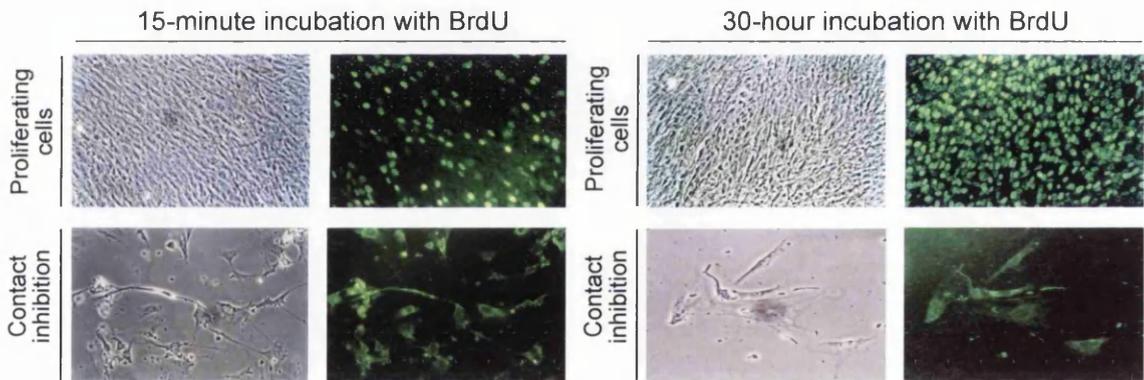


**Figure 7.2. Trinucleotide repeat dynamics in D4393K cells arrested by contact inhibition.** (A) The autoradiograph shown represent SP-PCR analyses of DNA samples extracted from D4393K cells arrested by contact inhibition. Fresh medium was added to the cultures every two or three days. Two replicate cultures, were analysed at each time point. Control cells were maintained for similar periods of time, and passaged when confluent. The number of population doublings (PD) is shown above each control panel. Twenty to 50 transgene molecules were amplified in each reaction. The molecular weight markers, converted into CTG repeat numbers, are displayed on the right. Note the expansion of a population of cells carrying longer repeats in both control and arrested cultures. Black arrowheads (◄) indicate evidence of clonal expansion. (B) The boxplots show the degree of repeat size variation in control, arrested and progenitor (P) cultures. The top and bottom of the boxes correspond to the third (Q3) and first quartiles (Q1), respectively, and the line across the box displays the median. The lines extending from the top and the bottom of the boxes, include values that fall inside the lower and upper limits:  $Q1 - 1.5(Q3 - Q1)$  and  $Q3 + 1.5(Q3 - Q1)$ . The hash symbols (#) denote the exclusion of longer alleles (>240 repeats) from the analysis. (C) The boxplots display the degree of repeat length variation within the subpopulations of cells carrying longer trinucleotide tracts.



**Figure 7.3. Immunocytochemical characterisation of D4393K kidney cells.**

The two pictures represent the immunostaining of D4393K kidney. Vimentin filaments were detected with a primary mouse monoclonal antibody raised against human vimentin, and a secondary anti-mouse IgM-FITC conjugate. All cells stained positive for vimentin, but negative for a panel of cytokeratins



**Figure 7.4. BrdU incorporation analysis on D4393K cells arrested by contact inhibition.**

BrdU incorporation was analysed in D4393K cells, to assess DNA synthesis in cultures arrested by contact inhibition. Confluent cultures, maintained on six-well plates for 123 days, were incubated with BrdU for 15 minutes or 30 hours, digested with trypsin and plated on an eight-well chamber slide. BrdU incorporated into genomic DNA was detected by immunostaining. Proliferating D4393K cells, grown and passaged as usual, were used as a positive control. Light micrographs are shown on the left, whereas immunodetection of BrdU is shown on the right. Note that most of the cells arrested by contact inhibition exhibited non-specific staining in the cytoplasm, but did not reveal nuclear staining, in contrast with proliferating cells.

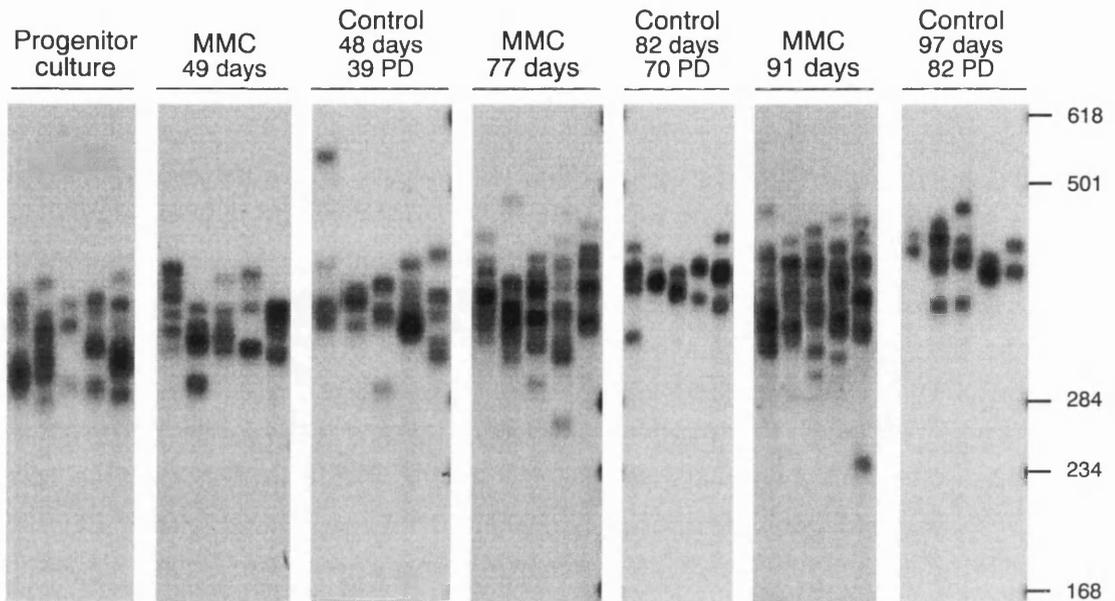
BrdU incubation, BrdU staining was not detected in the nucleus of cells arrested by contact inhibition, in clear contrast with the extensive nuclear BrdU staining observed in proliferating cells (Figure 7.4). An average of ~20-30% of the cells showed nuclear staining following a 15-minute incubation, whereas ~90% of the cell population displayed nuclear BrdU incorporation under proliferating conditions (Figure 7.4).

In summary, although BrdU incorporation and immunodetection analysis failed to reveal DNA synthesis in the nuclei of D4393K cells arrested by contact inhibition, the accumulation of longer alleles most likely resulted from the proliferation and clonal expansion of a few mutant cells, which were resistant to contact inhibition and happened to carry longer repeats than the average repeat number. Alternatively, in those cultures where repeat size variability was detected amongst the new mutants (Figure 7.2, 123 days of contact inhibition), the expanded alleles might have resulted from cell proliferation into the gaps left by dead cells, which tend to dissociate from the substrate, possibly allowing their neighbouring cells to divide.

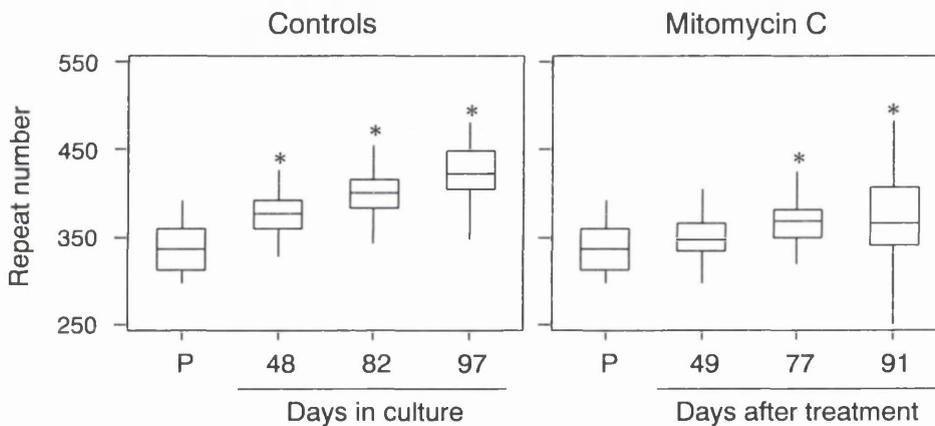
### **7.2.3. Trinucleotide repeat dynamics in Dmt-D cells arrested by mitomycin C**

As an alternative attempt to stop cellular proliferation in culture, spontaneously immortalised cell lines, in which the trinucleotide repeat dynamics had been previously described, were treated with mitomycin C. Replicate cultures derived from the D2763Kc2 clonal cell line, carrying a rapidly expanding trinucleotide repeat tract were treated with 30  $\mu$ M mitomycin C for three hours, and maintained in standard growth medium thereafter. Mitomycin C exposure resulted in the death of ~50% of the cells, as revealed by the levels of cell viability assessed by the acridine orange and ethidium bromide method (Section 2.4.6), one week after the treatment. However, no further decrease in cell viability was detected 49, 77 and 91 days following mitomycin C exposure, with the total number of living cells remaining unaltered. The repeat size variability was monitored at different time points by SP-PCR techniques, not only in cells exposed to mitomycin C, but also in control cells, which were maintained under proliferating conditions. High DNA input SP-PCR analyses revealed clear differences between the mutation profiles in mitomycin C treated cultures and progenitor cells (Figure 7.5.A). Twenty to 50 transgene molecules collected from each control and arrested cultures were accurately sized (Figure 7.5.B), and the analysis revealed that not only the control cells exhibited a significant median repeat gain of 0.897 units per day ( $p < 0.0001$ , two-tailed Mann-Whitney  $U$  test), but also mitomycin C treated cells, showed a significant median expansion of 0.341 repeats per day ( $p = 0.0389$ , two-tailed Mann-Whitney  $U$  test). One could argue that the differences observed in the median repeat number between mitomycin C treated cells and the progenitor culture are intimately associated with a selection for resistant cells, given the cell death caused by exposure to this DNA crosslinker agent. If that were the case, the median repeat size would be expected to stabilise as long as the total cell numbers remained constant. However, statistical analysis revealed significant differences in the median repeat length carried by mitomycin C treated cells between days 49 and 91 ( $p = 0.0484$ , two-tailed Mann-Whitney  $U$  test), a

A.



B.



**Figure 7.5. Dynamics of expanded CAG-CTG trinucleotide repeats in D2763Kc2 kidney cells arrested by mitomycin C.**

(A) The autoradiographs show representative SP-PCR analyses of DNA samples extracted from D2763Kc2 kidney cell cultures arrested following a three-hour exposure to 30  $\mu$ M mitomycin C (MMC) in standard growth medium. Treated cells were subsequently kept in standard culture medium without mitomycin C, and fresh medium was added every two to three days. Control cells were maintained in culture for the same period of time, and passaged when confluent. The number of population doublings (PD) for each control sample is displayed above each panel. Ten to 40 transgene molecules were amplified in each reaction. The molecular weight markers were converted into CTG repeat numbers and are shown on the right. (B) The boxplots show the degree of trinucleotide repeat length variability detected in control and mitomycin C arrested cells, as described in Figure 7.3. The median repeat size in both control and mitomycin C exposed cultures was compared with the median repeat number in the progenitor (P). Significant differences ( $p < 0.05$ , two tailed Mann-Whitney  $U$  test) are identified by asterisks (\*).

period when a decrease in the number of viable cells was not detected (Figure 7.5.B). Moreover, in addition to an increase in the median repeat length, the overall range of allele sizes also showed a shift towards longer repeat numbers, suggesting the appearance of new mutants in mitomycin C treated cultures.

BrdU incorporation and immunodetection techniques revealed very low levels of nuclear staining in mitomycin C treated cells per culture, even following a 30-hour incubation with BrdU (<5%) (Figure 7.6.A), indicative of some DNA synthesis, which in some cases is associated with distinctive nuclear foci (Figure 7.6.B). In contrast, ~90-95% of the proliferating cells showed positive labelling, following an identical incubation period.

The dynamics of expanded CAG•CTG repeats was also studied in two additional *Dmt-D* kidney cell lines, D2967K and D3111K (Table 3.2), following cell cycle arrest by mitomycin C exposure, and SP-PCR analysis as described above. Both cell lines carried slowly expanding trinucleotide sequences, and therefore bulk DNA SP-PCR amplifications might have masked subtle differences in allele size heterogeneity between mitomycin C treated cells and the progenitor culture. Single molecule analysis was then performed, and at least 100 transgene molecules collected from each culture were individually sized. Indeed, the repeat distributions confirmed very low levels of somatic instability in the control cells of both *Dmt-D* kidney lines, and an expansion bias was not always observed (Figures 7.7 and 7.8). However, some statistically significant differences were revealed, when the repeat distributions were compared between control cells and their progenitor cultures ( $p < 0.05$ , two-tailed Mann-Whitney *U* test). More interestingly, similar differences were also observed between some cultures arrested by mitomycin C, and the original cells from which all cultures were derived (Figures 7.7 and 7.8). To account for the hypothesis of mitomycin C-induced cell selection and preferential survival, as described previously, the repeat profiles in treated cultures were also compared to those detected in treated cells collected 41-44 days after cell cycle arrest, and significant differences were still detected (Figures 7.7 and 7.8).

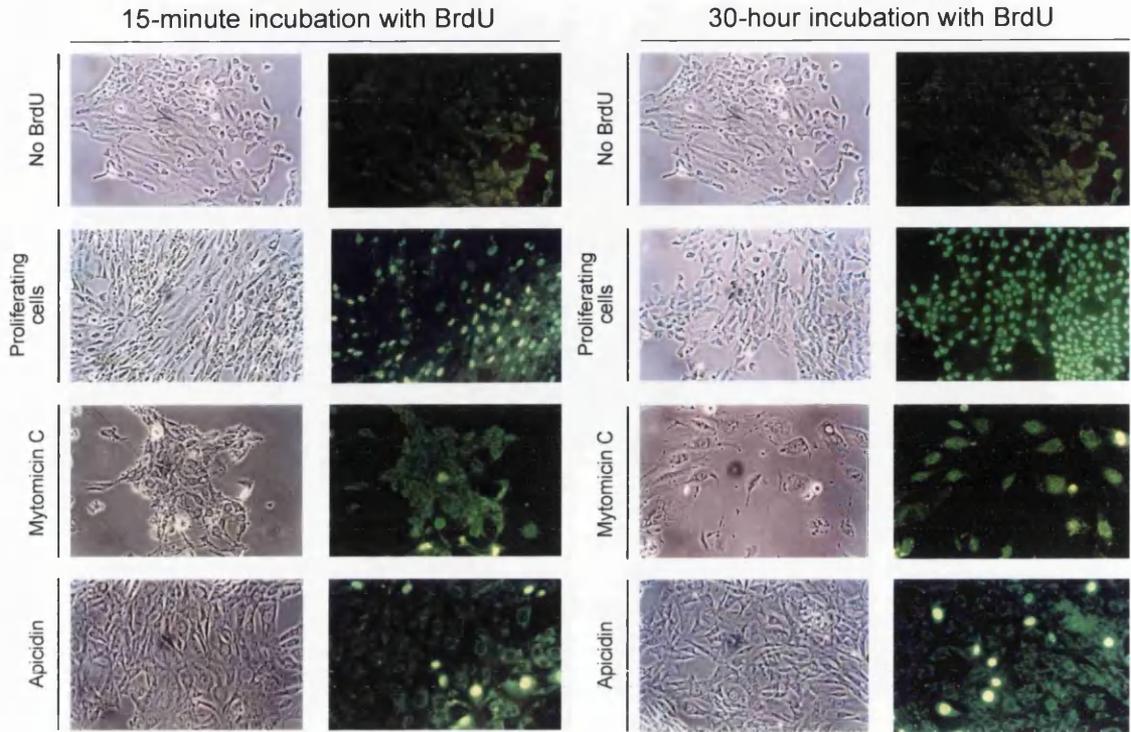
Although the results presented here, particularly in regards to D2967K and D3111K cells lines are not totally conclusive, and do not represent the ultimate experimental evidence for the accumulation of repeat size mutations in cells arrested by mitomycin C, they do suggest that variation in trinucleotide repeat size may happen independently of cell division and DNA replication.

#### **7.2.4. Trinucleotide repeat dynamics in *Dmt-D* cells arrested by apicidin**

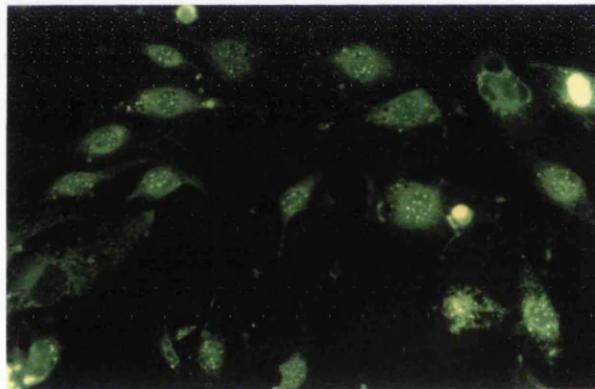
Although mitomycin C was able to inhibit cell proliferation of cultured *Dmt-D* cells, this chemical acts as a potent DNA crosslinker and is likely to affect multiple aspects of DNA metabolism apart from replication, including transcription and repair. Therefore, alternative ways of arresting cell proliferation in culture were explored.

Continuous exposure of six replicate D2763Kc2 and D979K cultures to 320 nM apicidin for a week resulted in a minor decrease (~9%) in the number of viable cells, as assessed by the

A.

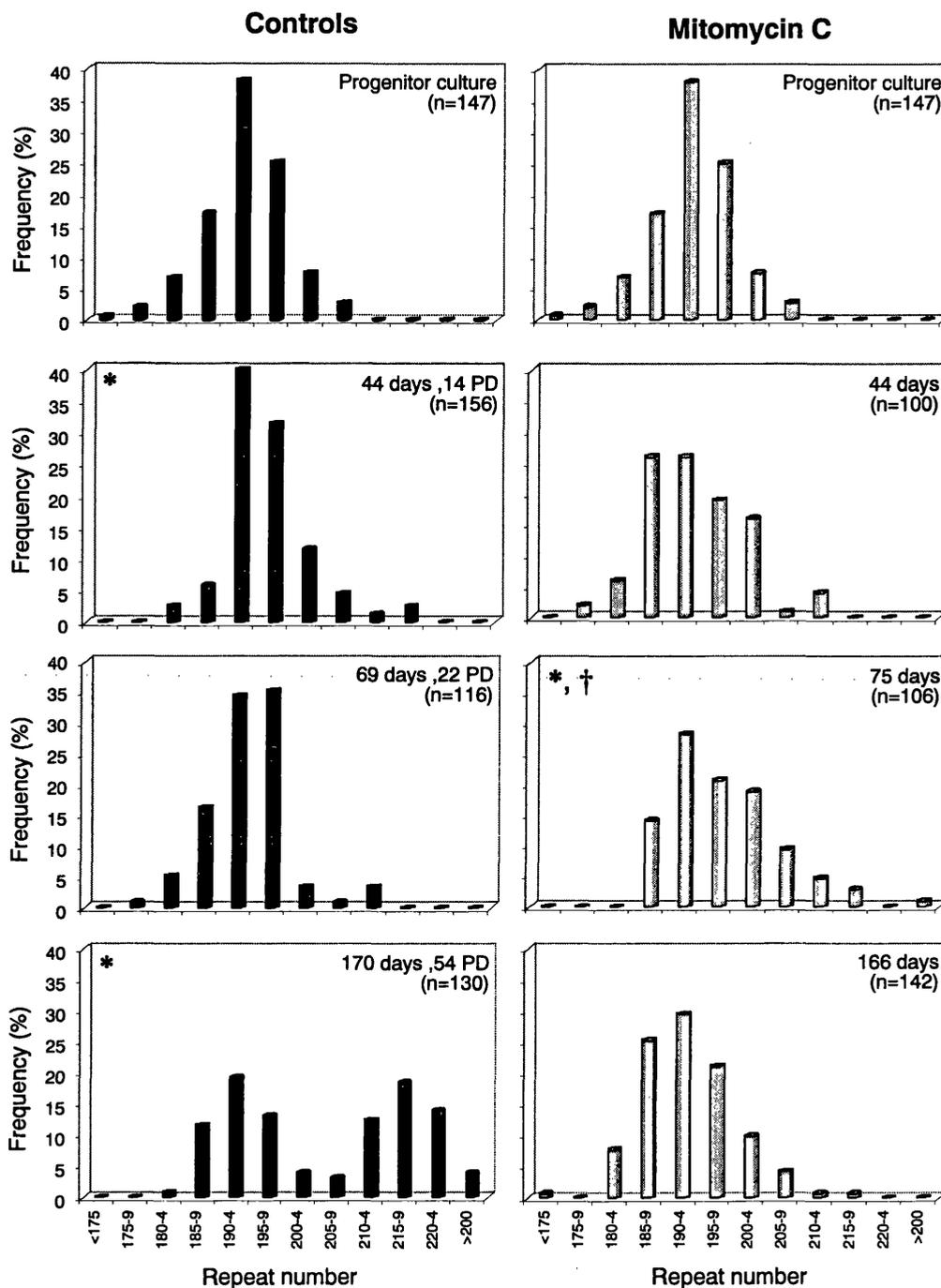


B.

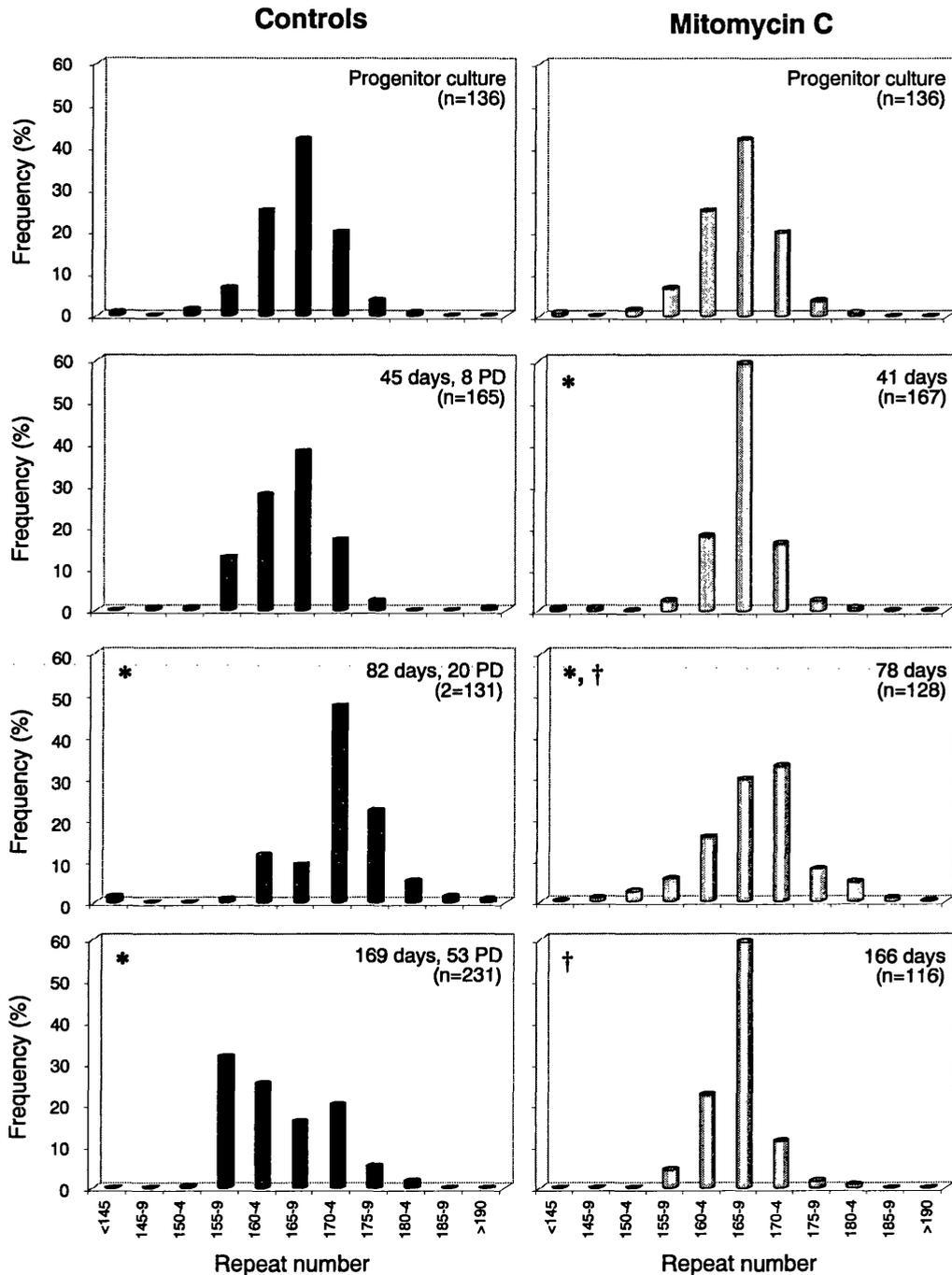


**Figure 7.6. BrdU incorporation analysis on D2763Kc2 kidney cells arrested by mitomycin C or apicidin exposure.**

(A) BrdU incorporation was analysed by immunocytochemistry techniques in D2763Kc2 kidney cells, in order to assess DNA synthesis in cells treated with mitomycin C or apicidin. D2763Kc2 cells, plated on eight-well chamber slides, were exposed to 30  $\mu$ M mitomycin C for three hours (as described in Figure 7.5), and the BrdU incorporation levels assayed one week later, following a 15-minute or a 30-hour exposure to the chemical. Alternatively, D2763Kc2 cells were continuously exposed to 320 nM apicidin for a week, and then stained for BrdU incorporation. The left columns show light micrographs, whereas immunodetection of BrdU is shown on the right. Non-specific staining was revealed in cells that were not incubated with BrdU. Proliferating D2763Kc2 kidney cells were used as a positive control. Note that most of the arrested cells exhibited non-specific cytoplasmic staining, but revealed considerably lower levels of bright nuclear staining, when compared to proliferating cells. (B) Also note the distinctive pattern of speckled nuclear staining in mitomycin C treated cells, following a 30-hour incubation with BrdU.



**Figure 7.7. Repeat distributions in D2967K cells arrested by mitomycin C.** D2967K kidney cells were treated with 30  $\mu$ M mitomycin C for three hours, and maintained in standard growth medium thereafter. Fresh medium without mitomycin C was added to the cultures every two to three days. Control cells were maintained under normal proliferating conditions, and the number of population doublings (PD) is displayed in each panel. DNA samples were collected at different time points and trinucleotide repeat profiles determined by single molecule SP-PCR analysis. For each time point at least 100 transgene molecules (n) were amplified by SP-PCR at low DNA concentration (1-3 DNA molecules per reaction) and individually sized. The median repeat number in each culture was compared with the progenitor cells, and significant differences ( $p < 0.05$ , two-tailed Mann-Whitney *U* test) are identified by asterisks (\*). Significant differences between treated cultures collected at later stages and cells collected 44 days after mitomycin C exposure are also indicated (†).



**Figure 7.8. Repeat distributions in D3111K cells arrested by mitomycin C.**

D3111K kidney cells were treated with 30  $\mu$ M mitomycin C for three hours, and maintained in standard growth medium thereafter. Fresh medium without mitomycin C was added to the cultures every two to three days. Control cells were maintained under normal proliferating conditions, and the number of population doublings (PD) is displayed in each panel. DNA samples were collected at different time points and trinucleotide repeat profiles determined by single molecule SP-PCR analysis. For each time point at least 100 transgene molecules (n) were amplified by SP-PCR at low DNA concentration (1-3 DNA molecules per reaction) and individually sized. The median repeat number in each culture was compared with the progenitor cells, and significant differences ( $p < 0.05$ , two-tailed Mann-Whitney *U* test) are identified by asterisks (\*). Significant differences between treated cultures collected at later stages and cells collected 41 days following mitomycin C exposure are also indicated (†).

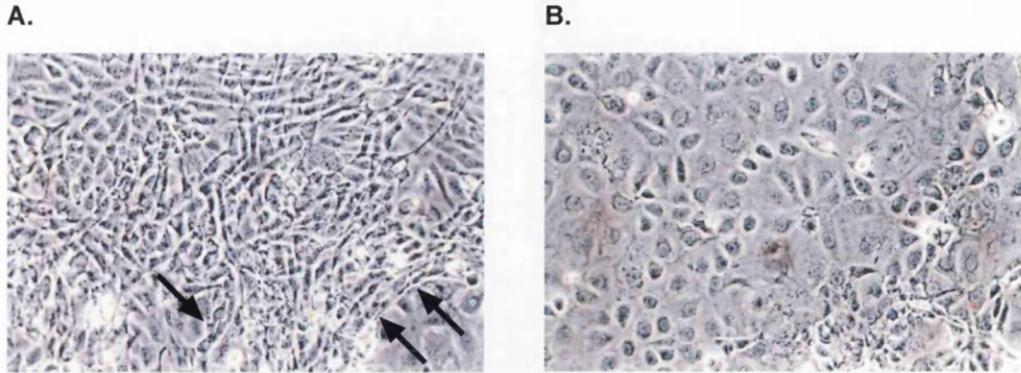
acridine orange and ethidium bromide method (Section 2.4.6). However, the difference did not prove to be statistically significant ( $p > 0.05$ , two-tailed  $t$  test). The total number of living cells appeared to remain constant one month later, when cell viability was assessed again by the same method. Multiple cell layers could not be observed by phase contrast microscopy, indicating that cells were not proliferating upon exposure to apicidin. The effects of apicidin on cell proliferation were associated with morphological changes in cell shape and cell-to-cell adhesion, suggesting that growth of *Dmt-D* kidney cells in apicidin restores contact inhibition (Figure 7.9). To confirm cell cycle arrest, BrdU incorporation and detection analyses were performed on both cell lines following one-week exposure to apicidin. Both D2763Kc2 and D979K treated cells exhibited very low levels of nuclear staining per culture (~5%-10%), indicative of low levels of DNA synthesis in the presence of this drug (Figures 7.6 and 7.10).

Trinucleotide repeat length variability was assessed in two treated replicate cultures at different time points, and compared with control cells, proliferating under normal growth conditions for similar periods of time. High DNA input SP-PCR analyses revealed that D2763Kc2 treated cultures not only continued to accumulate expanded alleles, but they were also exhibited a higher expansion rate than control cells (Figure 7.11.A). Single molecule analysis was performed to confirm these observations, and accurate sizing of 20 to 50 transgene molecules collected from each culture revealed that the median repeat number in apicidin treated cultures was in fact significantly higher than in the progenitor cells ( $p < 0.0003$ , two-tailed Mann-Whitney  $U$  test) at all time points analysed (Figure 7.11.B). More interestingly, while the median repeat number showed an increase of 0.426 units per day in the control cells, the apicidin treated cultures exhibited a significantly higher median repeat gain: 1.112 and 1.448 repeats per day ( $p < 0.0001$ , two-tailed Mann-Whitney  $U$  test), at 34 and 67 days of continuous apicidin exposure, respectively.

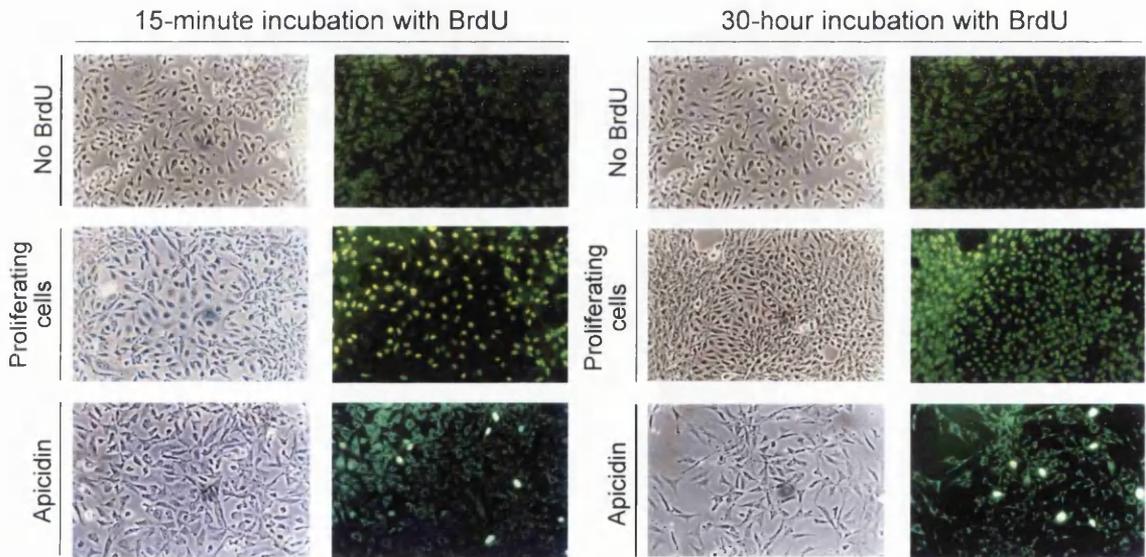
A similar analysis was performed on D979K cells arrested by continuous exposure to 320 nM apicidin. However, the analysis proved more difficult, not only because the progenitor culture consisted of two subpopulations of cells as revealed by SP-PCR amplification, but also because the subpopulation carrying longer alleles soon overgrew the control cultures, with the concomitant loss of cells carrying shorter repeats (Figure 7.12.A). Nevertheless, accurate sizing of 20 to 50 molecules revealed statistically significant differences, between some D979K cultures exposed to apicidin for 98 days and the progenitor culture ( $p < 0.05$ , two-tailed Mann-Whitney  $U$  test) (Figure 7.12.B).

As observed with mitomycin C treated cells, the increase in the median repeat size in apicidin treated cells, was accompanied by an expansion-biased trend in the repeat size range detected in both D2763Kc2 and D979K cultures (Figures 7.11 and 7.12). This finding strongly suggests that changes in the mutation profiles are indeed the result of the occurrence of new mutants in culture exposed to apicidin, rather than a simple cell selection process *in vitro*.

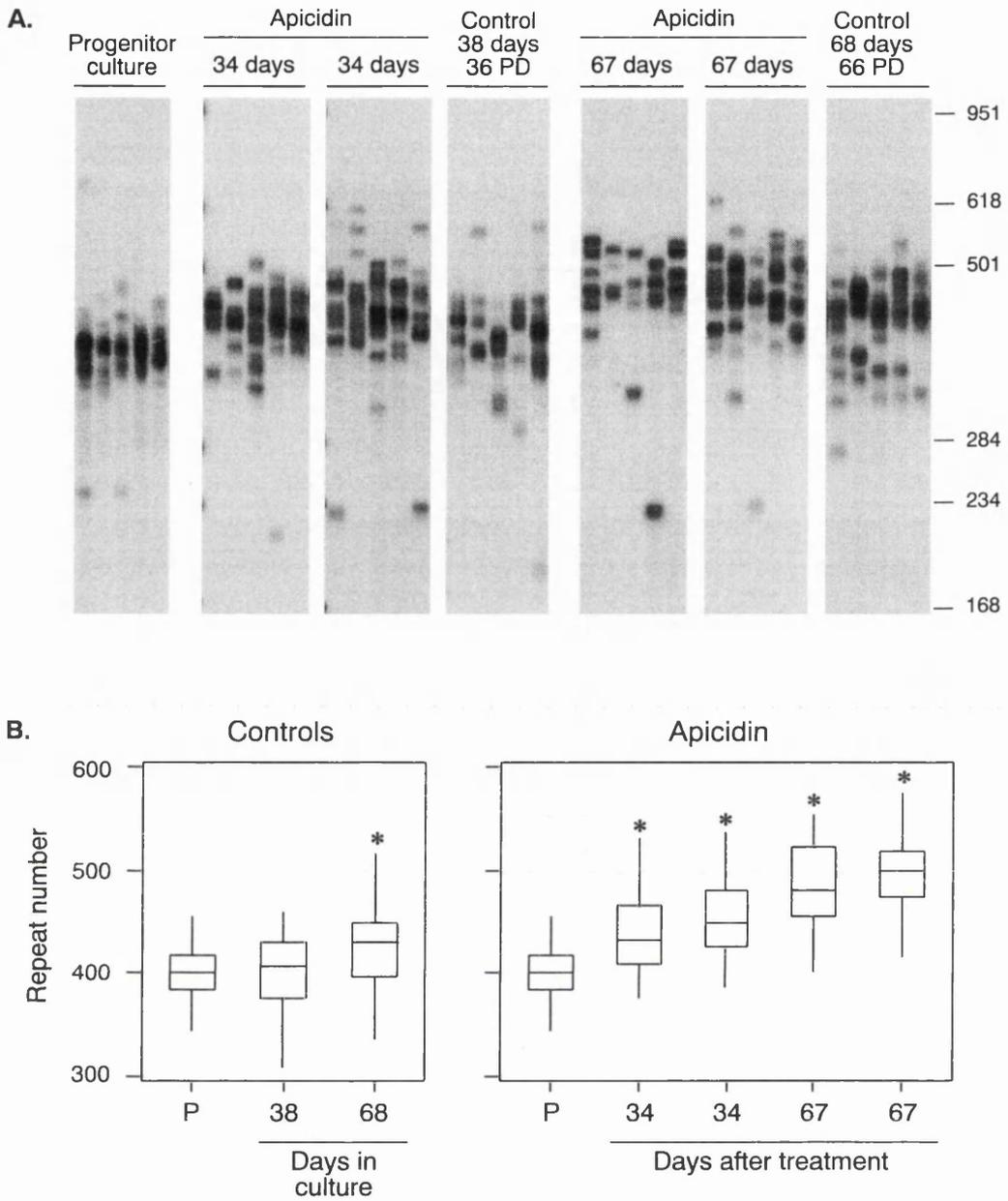
In summary, regardless of the extremely low levels of nuclear DNA synthesis detected in cells exposed to apicidin, the repeats carried by treated cultures exhibited trinucleotide repeat instability, suggesting that cell division is not strictly required for repeat length mutations to occur within expanded CAG•CTG.



**Figure 7.9. Morphological changes in *Dmt-D* kidney cells cultured in 320 nM apicidin.** The pictures represent phase-contrast microphotographs of D979K cells maintained in (A) complete growth medium or in (B) complete medium containing 320 nM apicidin. Note the occurrence of multiple cell layers in the overgrown culture (A), and the presence of areas of cells detached from the substrate (black arrows). In cultures exposed to apicidin, flattened cells form a uniform confluent monolayer (B).

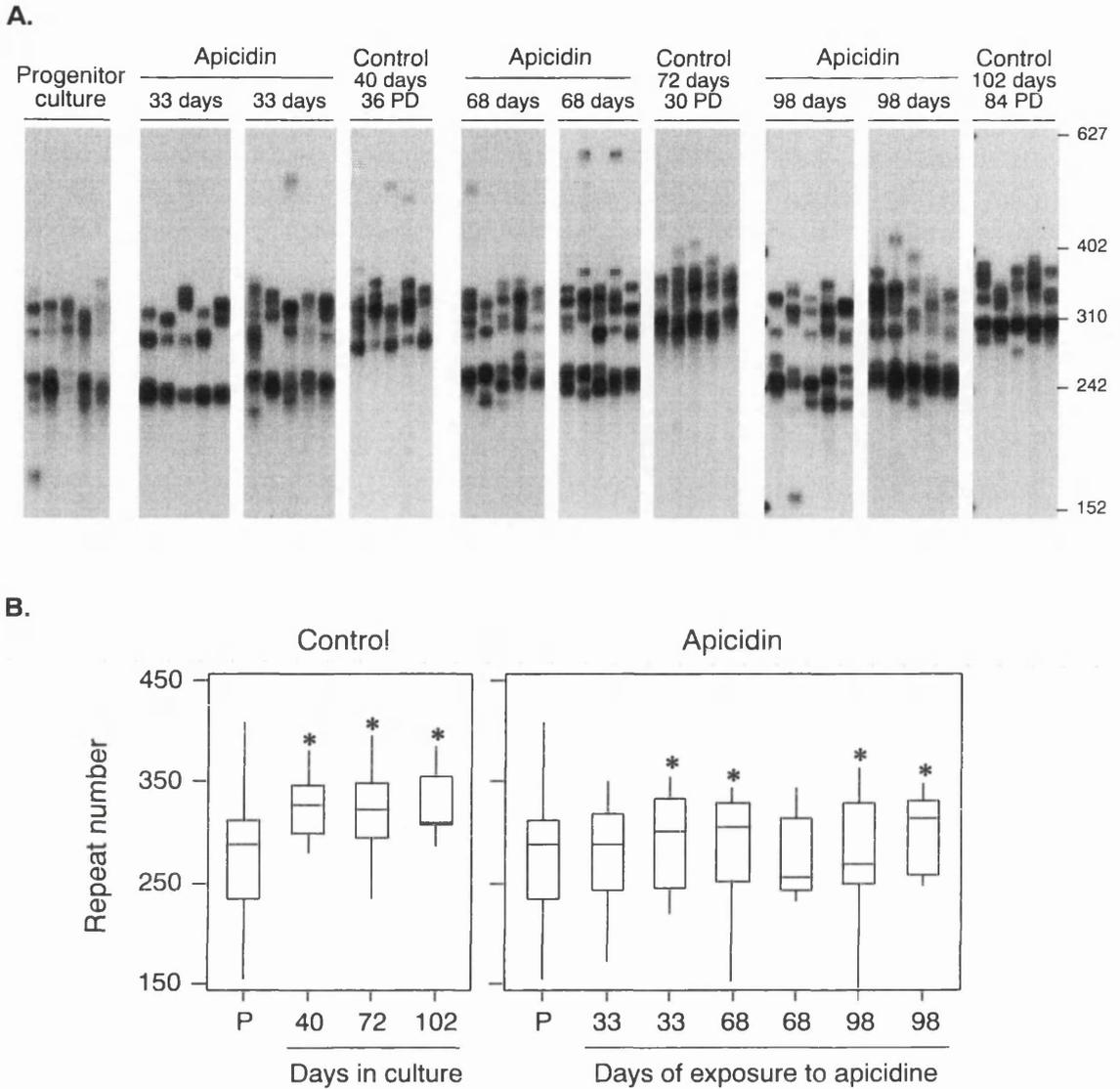


**Figure 7.10. BrdU incorporation analysis on D979K cells arrested by apicidin.** BrdU incorporation by D979K kidney cells was analysed by immunocytochemistry techniques, in order to assess DNA synthesis in cells treated with apicidin. D979K cells, plated on eight-well chamber slides, were exposed to 320 nM apicidin and the BrdU incorporation levels assayed one week later, following a 15-minute or a 30-hour exposure to the chemical. Light micrographs are shown on the left, whereas immunodetection of BrdU is shown on the right. Non-specific staining was revealed in cells that were not incubated with BrdU. Proliferating D979K cells were used as a positive control. Most of the arrested cells exhibited non-specific cytoplasmic staining in the cytoplasm, and only a few revealed nuclear staining.



**Figure 7.11. Dynamics of expanded CAG·CTG trinucleotide repeats in D2763Kc2 arrested by apicidin.**

(A) The autoradiographs show representative SP-PCR analyses of DNA samples extracted from D2763Kc2 kidney cell cultures arrested by continuous exposure to 320 nM apicidin. Fresh medium supplemented with the drug was added to the cultures every two to three days. Two replicate apicidin treated cultures were analysed at each time point. Control cells were maintained in culture for the same period of time, and passaged when confluent. The number of population doublings (PD) for each control sample is displayed above each panel. Ten to 30 transgene molecules were amplified in each reaction. The molecular weight markers were converted into CTG repeat numbers and are shown on the right. (B) The boxplots show the degree of trinucleotide repeat length variability detected in control and arrested cells, as described in Figure 7.3. The median repeat size in both control and apicidin exposed cultures was compared with the median repeat number in the progenitor culture (P). Significant differences ( $p < 0.05$ , two tailed Mann-Whitney  $U$  test) are identified by asterisks (\*).



**Figure 7.12. Dynamics of expanded CAG-CTG trinucleotide repeats in D979K cells arrested by apicidin.**

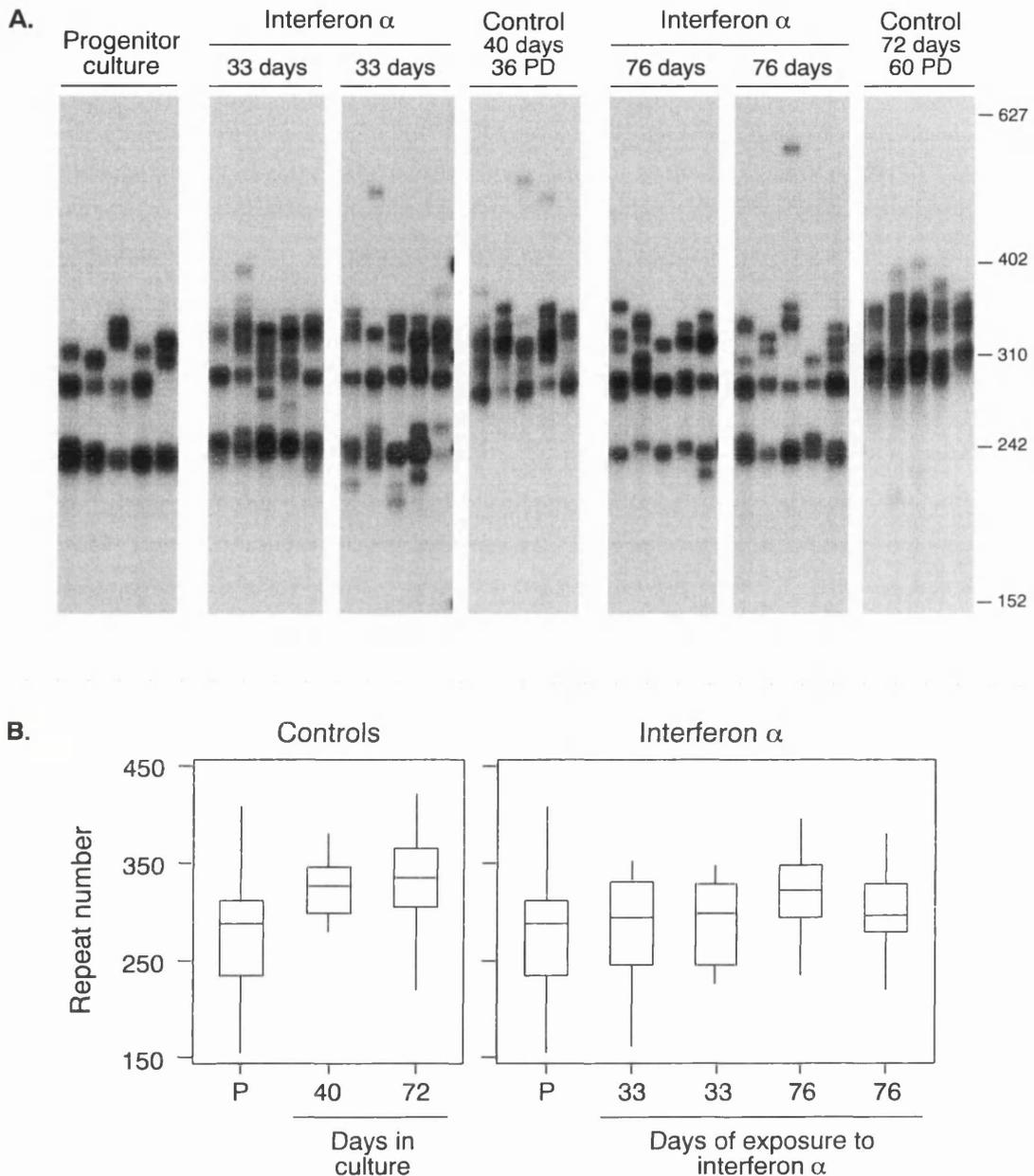
(A) The autoradiographs show representative SP-PCR analyses of DNA samples extracted from D979 kidney cell cultures arrested by continuous exposure to 320 nM apicidin. Fresh medium supplemented with the drug was added to the cultures every two to three days. Two replicate apicidin treated cultures were analysed at each time point. Control cells were maintained in culture for the same period of time, and passaged when confluent. The number of population doublings (PD) for each control sample is displayed above each panel. Ten to 40 transgene molecules were amplified in each reaction. The molecular weight markers were converted into CTG repeat numbers and are shown on the right. (B) The boxplots show the degree of trinucleotide repeat length variability detected in control and apicidin arrested cells, as described in Figure 7.3. The median repeat size in both control and apicidin exposed cultures was compared with the median repeat number in the progenitor culture (P). Significant differences ( $p < 0.05$ , two tailed Mann-Whitney  $U$  test) are identified by asterisks (\*).

### **7.2.5. Trinucleotide repeat dynamics in confluent Dmt-D cell cultures exposed to interferon $\alpha$**

Interferon  $\alpha$  was also used given its ability to arrest mammalian cell cycle progression. Both D2763Kc2 and D979K cells were continuously exposed to 500 U/ml interferon  $\alpha$ . Interferon  $\alpha$  failed to inhibit D2763Kc2 cell proliferation. D2763Kc2 cells proliferated continuously even in the presence of the polypeptide, eventually dying, possibly due to the lack of nutrients in the medium. Proliferation of D979K cells was initially inhibited by interferon  $\alpha$ , as revealed by their flattened morphology and absence of multiple cell layers. Following this initial period of apparent cell cycle arrest, D979K cells regained their proliferative capacity, as revealed by the growth of multiple cell layers, and rapid use of the nutrients, with simultaneous development of a characteristic yellow/orange colour in the culture medium. Nevertheless, these cells were capable of surviving under conditions of extreme cell density, as long as fresh medium, supplemented with the cytokine, was added to the cultures every two to three days. The treatment was therefore continued in order to investigate the effects of high cell density on the trinucleotide repeat dynamics in this cell line. Two replicate D979K cultures treated with interferon  $\alpha$  were collected at different time points, and the repeat size variability assessed by SP-PCR techniques. As discussed before, the analysis was complicated by the occurrence of a selective sweep within the control cell culture (Figure 7.13.A). Accurate sizing of 20 to 50 transgene molecules collected from each culture allowed the determination of the median repeat sizes and rate of expansion in treated and control cells (Figure 7.13.B). Not surprisingly, statistical analysis revealed significant differences between the two control cultures, and the progenitor culture ( $p < 0.0001$ , two-tailed Mann-Whitney  $U$  test), mainly due to the selection of cells carrying longer repeat tracts that happened over the first 40 days in culture. Most importantly, significant differences were also detected between the median repeat size in cells treated with interferon  $\alpha$  and in the progenitor culture ( $p < 0.01$ , two-tailed Mann-Whitney  $U$  test), suggesting that trinucleotide repeats carried by cells cultured at high density are still unstable, in spite of the low rate of cell turnover. Furthermore, a higher degree of repeat size variation, within both subpopulations, appears to occur in treated cells, compared to the progenitor culture. An extensive single molecule analysis, performed with a higher number of replicates, would be required to establish a possible difference, if any, between trinucleotide repeat expansion rates in confluent cells relative to proliferating cells.

## **7.3. Discussion**

To date, most discussion of the processes responsible for the dynamic behaviour of triplet repeat mutations has focused on replication-associated mechanisms such as DNA polymerase slippage (Richards and Sutherland, 1994). However, multiple lines of evidence support a mutation mechanism that may not be strictly dependent on cell division. To address the divergence between the assumptions resulting from the replication slippage mechanism, and the experimental data,



**Figure 7.13. Dynamics of expanded CAG-CTG trinucleotide repeats in D979K cells arrested by interferon  $\alpha$ .**

(A) The autoradiographs show representative SP-PCR analyses of DNA samples extracted from D979 kidney cell cultures arrested by exposure to 500 U/ml interferon  $\alpha$ . Fresh medium supplemented with interferon  $\alpha$  was added to the cultures every two to three days. Two replicate treated cultures were analysed at each time point. Control cells were maintained in culture for the same period of time, and passaged when confluent. The number of population doublings (PD) for each control sample is displayed above each panel. Ten to 40 transgene molecules were amplified in each reaction. The molecular weight markers were converted into CTG repeat numbers and are shown on the right. (B) The boxplots show the degree of trinucleotide repeat length variability detected in control and arrested cells, as described in Figure 7.3. The median repeat sizes in all controls and cultures exposed to interferon  $\alpha$  were found to be significantly different from the repeat size in the progenitor (P) culture ( $p < 0.05$ , two tailed Mann-Whitney  $U$  test).

*Dmt-D* mouse kidney cell lines, exhibiting CAG•CTG instability in culture, were grown under conditions of high cell density or cell cycle arrest, and the repeat stability monitored over time.

Low serum levels and inhibition by cell-to-cell contact failed to completely arrest proliferation. Although these cells were not passaged for long periods of time, and were indeed growing under conditions of extreme cell density, SP-PCR analysis revealed the accumulation of mutant alleles that in some cases clearly resulted from clonal expansion of mutant cells. Presumably, these subpopulations of proliferating cells exhibited enhanced survival and/or proliferative capacity under stress conditions imposed by high levels of confluency. Suboptimal growth conditions may result in lack of nutrients and possibly oxidative stress, therefore imposing high selective pressure on the cultured cells. Increased spontaneous mutagenesis has been reported in cells cultured in low serum (Rossman and Goncharova, 1998). The same cells contained higher levels of oxidant species, suggesting that the elevated spontaneous mutagenesis resulted from enhanced levels of oxidative stress. The results presented here illustrate that *Dmt-D* kidney cells exhibit CAG•CTG repeat instability, even when proliferation is limited.

Similarly, slowly proliferating interferon  $\alpha$  treated cells accumulated repeat length mutations. A simple comparison between the rates of expansion in treated cells versus controls was not possible, given the occurrence of a selective sweep in the control culture, that resulted in the loss of a cell subpopulation, still present in the interferon  $\alpha$  exposed cells.

Mitomycin C clearly impaired DNA synthesis, as revealed by the low levels of BrdU incorporation exhibited by D2763Kc2 cells following treatment. Inhibition of DNA synthesis was not accompanied by great levels of cell death, as the cell viability levels never dropped below 50% at any of the time points analysed. Similarly, continuous exposure to apicidin also inhibited cell division in both D2763Kc2 and D979K cultures, as inferred from the reduced BrdU immunostaining in treated cells, with minor cell death. Cultures exposed to mitomycin C or apicidin did not show a continuous accumulation of BrdU stained nuclei, from a short 15-minute incubation period, up to a long 30-hour exposure. In contrast, an increase from ~20-30% up to ~90-95% was detected in proliferating cells. This difference might suggest that the nuclear staining detected in presumed arrested cultures, rather than representing cells undergoing S phase and active DNA replication, may instead be the result of unscheduled DNA synthesis, attempted DNA replication, or even DNA repair synthesis. Interestingly, mitomycin C treated cells showed a peculiar pattern of speckled nuclear staining. It is not unreasonable to speculate that these foci may be associated with futile cycles of DNA repair, activated in these cells in an attempt to remove permanent damage caused by mitomycin C-induced crosslinks (Kao *et al.*, 2001). Considered together, these results suggest that the vast majority of the cells that survived the treatments were indeed arrested.

Trinucleotide repeat tracts continued to expand in D2763Kc2 cells treated with mitomycin C, yet at lower rates than in the controls. Nevertheless, significant higher median repeat numbers were found in exposed cultures, relative to the progenitor cells. Although significant differences, relative to the progenitor cell population, were detected in some D2967K or D3111K cultures exposed to mitomycin C, both cell lines failed to reveal continuing repeat expansion, in contrast to

D2763Kc2 cells. The low repeat instability detected in D2967K and D3111K cells may be due to the inherent low somatic mosaicism detected in both cell lines, even under conditions of rapid cell proliferation. Regardless of the extent of repeat instability observed, these results indicate that cells arrested by mitomycin C can still accumulate repeat length mutations in the absence of cell division.

The response of mammalian cells to mitomycin C is multi-faceted, including inhibition of DNA synthesis, cell cycle arrest at S and G<sub>2</sub>/M phases and induction of apoptotic cell death (Kang *et al.*, 2001). The apoptotic index peaks shortly after exposure to mitomycin C, decreasing thereafter (Castaneda and Kinne, 1999; Clarke *et al.*, 1997; Kang *et al.*, 2001). It could be therefore suggested that the differences in the mutation profiles of treated cells between different time points were due to cell death, and preferential survival of cells carrying particular repeat lengths. If that were the case, a stabilisation of the CAG•CTG tract would be expected following the initial period of cell loss. However, not only the median repeat number continued to increase in exposed D2763Kc2 cells, but also greater expansion-biased repeat size ranges were observed over time. Moreover, significant differences in the median repeat number were found between D2967K and D3111K cell samples collected at later stages and the first cell sample, collected more than 40 days following mitomycin C treatment. The continuous repeat accumulation of repeat length variation strongly supports the occurrence of trinucleotide instability in arrested cells.

Although a great deal is known about the chemistry of mitomycin C attack on DNA, little is understood about the recognition and repair of interstrand crosslinks. Since these lesions involve both strands of duplex DNA they present special challenges to the repair machinery, being difficult to repair. No evidence of significant repair has been reported for mitomycin C crosslinks, suggesting these lesions can subsist in mammalian cells for extended periods of time (Larminat *et al.*, 1998). Nevertheless, a number of multi-step DNA repair pathways, including NER, homologous recombination, and cell cycle checkpoint proteins are activated and take part in the recognition and attempted repair of DNA interstrand crosslinks (Dronkert and Kanaar, 2001). NER proteins recognise mitomycin C induced DNA crosslinks, and participate in the recruitment and formation of multimeric complexes on the crosslinked DNA (de Laat *et al.*, 1999; Mustra *et al.*, 2001; Warren *et al.*, 1998; Wood, 1997). It may be speculated that the recruitment and activation of a series of repair proteins may affect the trinucleotide repeat tract, leading to alternative processing of intermediate structures, and ultimately resulting in repeat length mutation. It should also be noted that two distinct types of lesions are produced in response to mitomycin C through different pathways: bulky adducts and crosslinks, resulting from mono-alkylation and bis-alkylation of DNA (Iyer and Szybalski, 1963; Tomasz *et al.*, 1987); and DNA lesions produced by reactive oxygen species, generated during futile cycles of oxidation and reduction of mitomycin C (Doroshov, 1981; Pritsos and Sartorelli, 1986). Therefore, mitomycin C-generated oxygen radical species may themselves play a critical role in inducing repeat length mutations in arrested cells.

The analysis of CAG•CTG repeat dynamics in D2763Kc2 cells continuously exposed to apicidin was particularly interesting. The expanded repeat tract not only continued to expand in arrested cells, but also appeared to expand more rapidly in non-dividing conditions relative to the

proliferating controls. Although an extensive measurement of cell viability was not carried out at all time points, and a higher number of replicates would be required to reach greater levels of statistical significance, this result may imply that, under certain conditions, the mutation rate of expanded trinucleotide repeats is actually higher in non-dividing cells than in rapidly proliferating cultures. Interestingly, the mutator phenotype of some MMR-deficient cell lines was shown to depend on growth conditions, with enhanced accumulation of mutations within dinucleotide microsatellites in tumour cells when growth was limited (Richards *et al.*, 1997). These results raise the possibility that repeat size mutations may occur in a time-dependent manner, in the absence of cell proliferation.

The interpretation of the results obtained with D979K cells, following exposure to both apicidin and interferon  $\alpha$ , was again complicated by the occurrence of selective sweeps, and could only be resolved by a detailed single molecule analysis for each of the cell subpopulations present in some of the cultures. Nevertheless, monitoring of the repeat dynamics in D979K cells revealed the accumulation of repeat length variation in cultures exposed to both chemicals, again hinting at a mutation mechanism that is not necessarily dependent on cell division.

Inhibition of histone deacetylase by apicidin was associated with morphological changes in *Dmt-D* kidney cells, as previously reported (Han *et al.*, 2000). There is the possibility that the effects of apicidin treatment on *Dmt-D* kidney cultures may not be uniquely attributed to the histone deacetylase inhibitor, as apicidin was initially dissolved in DMSO, and a final concentration of 0.02% (v/v) DMSO was present in the growth medium. At a concentration of 1.5% (v/v), DMSO efficiently restores contact inhibition, and arrests Chinese hamster ovary cells in the G<sub>1</sub> stage as a confluent monolayer. The restored contact inhibition phenotype is accompanied by suppression of apoptotic cell death usually observed in high cell density cultures. The arrest in cell cycle progression and the inhibition of apoptosis in response to DMSO is associated with increased p27 levels, morphological changes in cell shape and cell-to-cell adhesion capacity (Fiore and Degrossi, 1999). Therefore, despite its low concentration in the medium, the possibility that DMSO might have also played a role in the inhibition of cell cycle progression must be considered.

Regardless of the actual cause of cell cycle arrest, the results presented here suggest that, under particular conditions, non-dividing cells can still accumulate trinucleotide repeat length mutations. The possibility that multiple rounds of DNA damage and repair may mediate mutation instability in non-dividing cells is corroborated by the observation that the absence of *Msh2* (Manley *et al.*, 1999b), *Msh3* (van Den Broek *et al.*, 2002) and *Pms2* (Chapter 8) MMR gene products appears to reduce the degree of repeat size variability of expanded CAG•CTG tracts in mouse models that replicate trinucleotide repeat somatic mosaicism. In fact, the specific MMR activities of extracts derived from the entirely post-mitotic cells of young and senescent *Drosophila* adults were similar to those extracts derived from rapidly dividing embryos, suggesting the MMR system is still functional in post-mitotic and senescent tissues, maintaining DNA sequence integrity even in non-dividing cells (Bhui-Kaur *et al.*, 1998). In addition, the findings described, particularly for apicidin, open new avenues to the study of the effect of genotoxic agents on the dynamics of expanded trinucleotide repeats. This potent antiproliferative histone deacetylase inhibitor appears

to present the potential to be used in association with other genotoxic agents, in order to test their ability to modify repeat dynamics, with the abolishment, or at least minimisation, of the effect of selective sweeps.

Interestingly, histone deacetylase inhibitors can rescue the pathological effects of toxic polyglutamine peptides *in vivo*. Both genetic and pharmacological reductions in the activity of histone deacetylase in transgenic *Drosophila* models expressing expanded polyglutamine tracts, either alone, or in the context of *HD* exon 1, retarded (or even arrested) neuronal degeneration (Steffan *et al.*, 2001). Potential therapeutical approaches based on histone deacetylase inhibitors were therefore proposed to prevent neurodegeneration, characteristic of polyglutamine disorders. Despite the short-term benefits resulting from the administration of histone deacetylase inhibitors (Steffan *et al.*, 2001), the results described in this chapter suggest that in the long-term, histone deacetylase inhibitors may result in an increased expansion rate, thereby possibly leading to a more severe symptomatology.

## 8. *Pms2* as a genetic modifier of trinucleotide repeat dynamics

### 8.1. Introduction

Instability within repetitive tracts of DNA is a feature shared by most, if not all organisms. The expansion of trinucleotide repeat sequences within the human genome has been implicated as the cause of an ever-increasing number of severe human conditions (Cummings and Zoghbi, 2000a; Cummings and Zoghbi, 2000b), and yet the molecular mechanisms that drive the expansion of these sequences remain unclear (Richards, 2001; Sinden, 2001).

In the attempt to develop an insight into the biology of triplet repeats, bacteria (Jaworski *et al.*, 1995; Kang *et al.*, 1995b; Sarkar *et al.*, 1998; Schumacher *et al.*, 1998) and yeast (Freudenreich *et al.*, 1997; Maurer *et al.*, 1996; Miret *et al.*, 1997) were initially used in the search for factors that influence repeat tract stability *in vivo*. Despite the simplicity of these model systems, these studies revealed a complex picture, in which the dynamics of trinucleotide repeat sequences is affected by multiple aspects of DNA metabolism, including DNA replication (Iyer and Wells, 1999; Kang *et al.*, 1995b; Kang *et al.*, 1996), transcription (Bowater *et al.*, 1997), and various pathways of DNA repair (Jaworski *et al.*, 1995; Parniewski *et al.*, 1999; Parniewski *et al.*, 2000; Rolfsmeier *et al.*, 2000; Schmidt *et al.*, 2000; Schumacher *et al.*, 1998; Schweitzer and Livingston, 1997). Different mutation mechanisms have been suggested, yet most of them only apply to simple model systems, where deletion-biased trinucleotide repeat instability appears to primarily depend on DNA replication. The most favoured mechanism, the DNA polymerase slippage model, predicts that trinucleotide repeat instability is dependent on cell division, and also that MMR gene mutations are associated with increased levels of repeat length variation, as a result of impaired repair of slipped-stranded DNA secondary structures. Although the analysis of trinucleotide repeat dynamics in microbial models has provided some support in favour of this mechanism (Wells, 1996), little is understood about the molecular pathways that control triplet repeat instability in a complex mammalian cell environment. Mouse models have been used as powerful tools to understand the dynamics of these mutations (Gourdon *et al.*, 1997; Lorenzetti *et al.*, 2000; Mangiarini *et al.*, 1996; Monckton *et al.*, 1997; Wheeler *et al.*, 1999b). Monitoring somatic mosaicism in tissues collected from these mice revealed the lack of correlation between levels of trinucleotide instability and the rates of cell turnover (Fortune *et al.*, 2000; Kennedy and Shelbourne, 2000; Lia *et al.*, 1998; Seznec *et al.*, 2000). Moreover, the levels of repeat instability in homogeneous cell lines do not correlate with cell division rates *in vitro* (Chapter 3). Taken together, these observations have raised serious reservations about the replication slippage model (Richards and Sutherland, 1994), which would predict greater levels of somatic mosaicism in rapidly proliferating tissues or cell lines. It is therefore conceivable that the expansion mechanism operating in a mammalian scenario is not entirely dependent on DNA replication, embracing a complex series of other major factors.

DNA repair, particularly the DNA mismatch repair (MMR) system, was soon considered as a plausible modifier of trinucleotide repeat instability, since it is required to maintain genomic integrity in both prokaryotes and eukaryotes and stabilise the cellular genome, by correcting single mismatches, and short unpaired regions within DNA, such as small insertions and deletions. The current mammalian model of MMR implicates heterodimers formed between MSH2 and either MSH3 or MSH6 (MutSB and MutS $\alpha$ , respectively) as being critical for the initial recognition of single base mismatches or small insertions and deletions. Recruitment of a second heterodimer made up of MLH1 and either PMS2 (MutL $\alpha$ ) or MLH3 (MutL $\beta$ ) is thought to be essential for subsequent excision and resynthesis, finally leading to DNA repair (Jiricny, 2000; Peltomaki, 2001a). Upon inactivation of MMR, increased heterogeneities are observed at simple repetitive DNA sequences, such as mono- and dinucleotide tracts (Buermeier *et al.*, 1999; Harfe and Jinks-Robertson, 2000; Jiricny, 2000). The association of impaired MMR activity and elevated genetic instability at simple DNA repeat sequences is particularly strong in hereditary non-polyposis colorectal cancer (HNPCC) (Peltomaki, 2001b; Toft and Arends, 1998). Unlike HNPCC, which is characterised by genetic instability at all simple DNA repeats throughout the genome, trinucleotide repeat diseases are associated with the expansion of a repetitive sequence within a single locus. In fact, no dinucleotide instability is observed in the genome of HD patients (Goellner *et al.*, 1997). Moreover, mutations in MMR genes are not solely responsible for triggering trinucleotide instability of normal sized CTG•CAG or CGG•CCG repeat tracts at the *DM1* or *FRAXA* loci, respectively, as assessed by bulk DNA PCR techniques (Kramer *et al.*, 1996). Taken together, these observations suggest that microsatellite dynamics might be controlled by distinct molecular mechanisms in triplet repeat disorders and in HNPCC. Nevertheless, MMR mutations may still affect trinucleotide repeat dynamics, and extensive analysis of the role of MMR on triplet repeat expansion has been carried out in bacteria and yeast.

MMR has been reported to play opposite roles in the stability of CAG•CTG repeat arrays cloned into *E. coli*. While mutations in MMR genes stabilise long (CAG•CTG)<sub>130-180</sub> repeat tracts (Jaworski *et al.*, 1995), the opposite result was observed with shorter (CAG•CTG)<sub>64</sub> sequences (Schumacher *et al.*, 1998). It was finally suggested that MMR might have two opposing effects on CTG•CAG repeat arrays: preventing single repeat unit insertions and deletions, and increasing the frequency of large contractions (Parniewski *et al.*, 2000; Schmidt *et al.*, 2000). MMR is also the major stabilising force of trinucleotide repeats in *S. cerevisiae* (Rolfsmeier *et al.*, 2000; Schweitzer and Livingston, 1997). Mutations in other repair genes, such as those involved in the processing of Okazaki fragments, have also been shown to destabilise trinucleotide repeat tracts, in both bacteria (Henricksen *et al.*, 2000) and yeast (Freudenreich *et al.*, 1998; Spiro *et al.*, 1999). The SOS system has also been proposed as a modifier of AC•TG dinucleotide instability in *E. coli* (Morel *et al.*, 1998).

In humans *PMS2* gene mutations have been associated with dramatic microsatellite instability in both normal and tumour tissues in Turcot syndrome, which is characterised by an association of malignant tumours of the central nervous system and colon cancer (De Rosa *et al.*, 2000; Miyaki *et al.*, 1997). *PMS2* mutations have also been detected in HNPCC kindreds

(Nicolaides *et al.*, 1994). In order to gain a better understanding of the function of the *Pms2* gene product in the MMR pathway of higher eukaryotes, *Pms2* knock-out mice have been generated by gene targeting procedures (Baker *et al.*, 1995). In these animals, exon 2 of the *Pms2* gene was replaced by a neomycin (*neo*) resistance cassette, resulting in total absence of *Pms2* protein in homozygotes (Baker *et al.*, 1995). *Pms2*-null homozygotes develop spontaneous lymphomas and sarcomas, but not intestinal tumours, whereas heterozygous mice for the *Pms2* disruption develop normally, with no predisposition to any tumours (Baker *et al.*, 1995; Prolla *et al.*, 1998). However, both heterozygous (Qin *et al.*, 2000) and homozygous (Qin *et al.*, 1999) *Pms2* knock-out mice are more susceptible to chemical carcinogenesis induced by alkylating agents than wild-type. *Pms2*-deficient mice are hypermutable to ionising radiation compared to wild-type animals, as revealed by sequence analysis of long mononucleotide repeat tracts (Xu *et al.*, 2001), suggesting that *Pms2*-deficient mice have an altered response to DNA damage (Xu *et al.*, 2001). In addition, lower levels of apoptosis have been reported in *Pms2*-deficient mouse cells, following exposure to ionising radiation (Zeng *et al.*, 2000), hinting at a role for the *Pms2* gene product in damage-induced programmed cell death.

*Pms2* nullizygous male mice are sterile, due to disrupted chromosome synapsis taking place during meiotic prophase I, providing evidence for the involvement of *Pms2* in meiotic recombination (Baker *et al.*, 1995).

Microsatellite analysis revealed an association between a *Pms2* null mutation and destabilised dinucleotide repetitive sequences in the male germline, in tumour and in tail DNA samples. Therefore, the *Pms2* gene product appears to play a role in DNA mismatch repair in multiple cell lineages (Baker *et al.*, 1995). Further studies revealed that *Pms2*-null homozygous mice have shown high levels of spontaneous microsatellite mutation (up to 100-fold above the wild-type background) in mononucleotide sequences in multiple tissues, even in those that are not associated with an increased risk of cancer, such as skin, liver, spleen, colon, brain and lung, confirming a key role for *Pms2* in the maintenance of genomic stability (Narayanan *et al.*, 1997). In addition, *Pms2* inactivation has a major impact on the frequency of deletions and insertions, with a significantly smaller effect on the occurrence of base substitutions (Narayanan *et al.*, 1997), which may be highly significant in the expansion mechanism of trinucleotide repeats. Analysis of the immunoglobulin gene variable regions in *Pms2*-deficient mice has also revealed an altered spectrum of mutation supporting a role of *Pms2* in somatic hypermutation (Casalho *et al.*, 1998; Winter and Gearhart, 2001; Winter *et al.*, 1998). Despite these high levels of spontaneous somatic mutations, *Pms2* disruption appears to be compatible with mostly normal development and life (Baker *et al.*, 1995), making these animals a powerful tool to study trinucleotide dynamics in a repair-deficient background.

In summary, mounting evidence suggests the involvement of *Pms2* in a wide range of biological processes, increasing the interest of investigating the impact of a *Pms2*-deficient genetic background on trinucleotide repeat length mutation.

In an attempt to clarify the possible role of MMR enzymes, specifically of MutL homologues, as modifiers of trinucleotide repeat dynamics in mammalian cells *in vivo*, *Dmt-D*

mice (Monckton *et al.*, 1997), which show great levels of tissue-specific age-dependent expansion-biased somatic mosaicism (Fortune *et al.*, 2000), were bred onto a MMR-deficient genetic background lacking *Pms2* (Baker *et al.*, 1995), and the levels of repeat instability assessed in different tissues for each of the different *Pms2* status.

## 8.2. Results

### 8.2.1. Breeding *Dmt-D* mice onto a *Pms2*-deficient genetic background

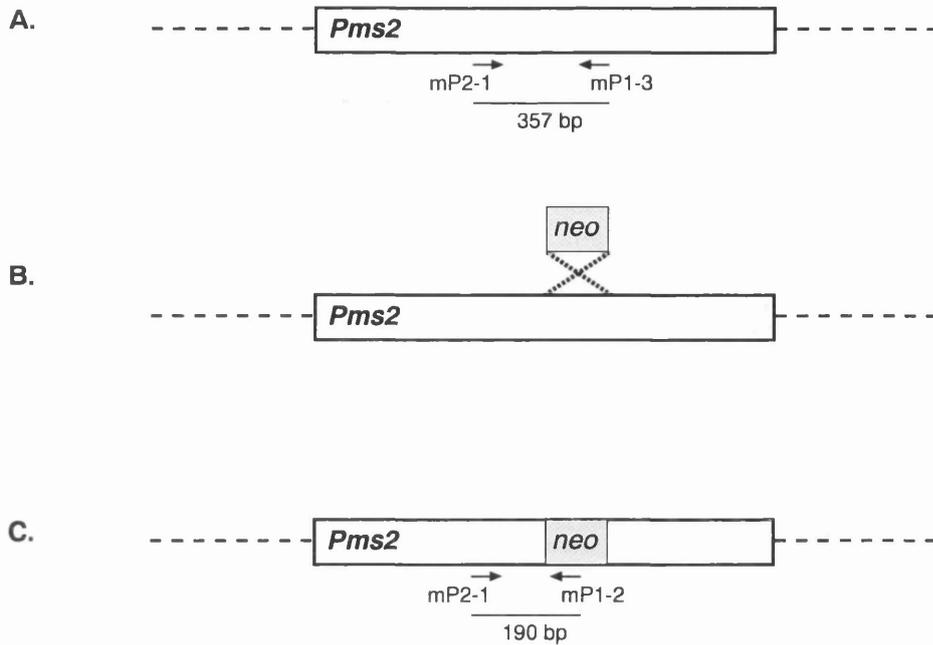
In order to breed *Dmt-D* mice onto a *Pms2*-deficient genetic background, mating pairs were initially set up between *Dmt-D* mice (hemizygous for the transgene) and heterozygous mice for the *Pms2* disruption (*Pms2*<sup>+/-</sup>). Mating pairs were then set up from subsequent generations between mice that carried both a *Dmt-D* transgene and a single *Pms2* gene disruption (*Dmt-D/Pms2*<sup>+/-</sup>), and mice that were only heterozygous for the *Pms2* mutation (*Pms2*<sup>+/-</sup>).<sup>5</sup>

### 8.2.2. Mouse genotyping for the *Pms2* deletion

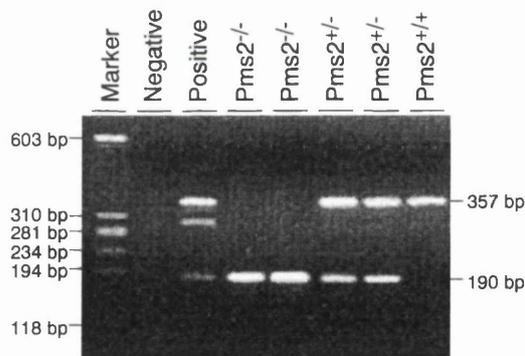
Following the crossing between *Dmt-D* and *Pms2*-deficient mice, a genotyping procedure was performed on the progeny in order to determine, not only the presence or absence of the *Dmt* transgene, but also to establish the *Pms2* status of each mouse. *Dmt* genotyping was carried out as described previously (Section 3.2.1). To determine the number of functional *Pms2* alleles carried by each mouse a PCR analysis was performed, using the oligonucleotide primer combination mP2-1, mP2-2 and mP2-3 (Table 2.7). The mP2-1 and mP2-3 primer set amplifies the undisrupted *Pms2* allele, generating a PCR product of 357 bp in size. Alternatively, a *Pms2* allele that has been disrupted by the insertion of the neomycin gene will give rise to a shorter PCR product (190 bp), by amplification with primers mP2-1 and mP2-2 (Figure 8.1). Therefore, a mixture of the three primers will allow the identification of mice carrying both functional *Pms2* alleles (*Pms2*<sup>+/+</sup>), mice heterozygous for the deletion (*Pms2*<sup>+/-</sup>) and nullizygous animals (*Pms2*<sup>-/-</sup>). Following amplification, the PCR products were resolved by agarose gel electrophoresis, and the mice genotypes determined according to the number and size of the PCR products observed for each reaction (Figure 8.2). The progenitor repeat length inherited by each mouse was calculated based on the analyses on tail DNA collected at weaning. Given the lack of trinucleotide instability at this stage, it is realistic to assume that the observed repeat size determined at this point, is a good estimate of the repeat size inherited by the mouse.<sup>6</sup>

<sup>5</sup> Mouse mating pairs were set up by Teresa Fortune (Division of Molecular Genetics, University of Glasgow, Glasgow, UK).

<sup>6</sup> Some of the genotyping analyses were performed by Teresa Fortune and John McAbney (Division of Molecular Genetics, University of Glasgow, Glasgow, UK).



**Figure 8.1** Mouse *Pms2* gene disruption by insertion of a *neo* resistance cassette. (A) Structure of mouse undisrupted *Pms2* gene showing the annealing sites of primers mP2-1 and mP2-3. (B) Disruption of mouse *Pms2* gene by the insertion of a *neo* resistance cassette by homologous recombination. (C) Disrupted *Pms2* gene, showing the annealing sites of primers mP2-1 and mP2-2 (Baker *et al.*, 1995)



**Figure 8.2.** Mouse *Pms2* genotyping by PCR analysis. Ten to 100 ng of mouse tail DNA were amplified in a 10  $\mu$ l PCR reaction with primers mP2-1, mP2-2 and mP2-3. The reactions were cycled 30 times, with an annealing temperature of 61°C. Following amplification, the PCR products were resolved in an ethidium bromide-stained 2% (w/v) agarose gel, and the mice genotypes determined according to the number and size of the PCR products observed for each reaction. Negative: PCR negative control with no template DNA added to the reaction. Positive: PCR positive control, using a DNA sample from a *Pms2*<sup>+/-</sup> mouse. Extra PCR products of a similar size to the 357-bp product of interest are observed in this lane. Lanes 4-8: PCR products generated by the amplification of tail DNA samples from *Pms2*<sup>+/+</sup>, *Pms2*<sup>+/-</sup> or *Pms2*<sup>-/-</sup> mice. The *Pms2* status for each animal is indicated above the lane. The sizes of the molecular weight markers are shown on the left.

### 8.2.3. Somatic mosaicism in heterozygous *Pms2* knock-out mice

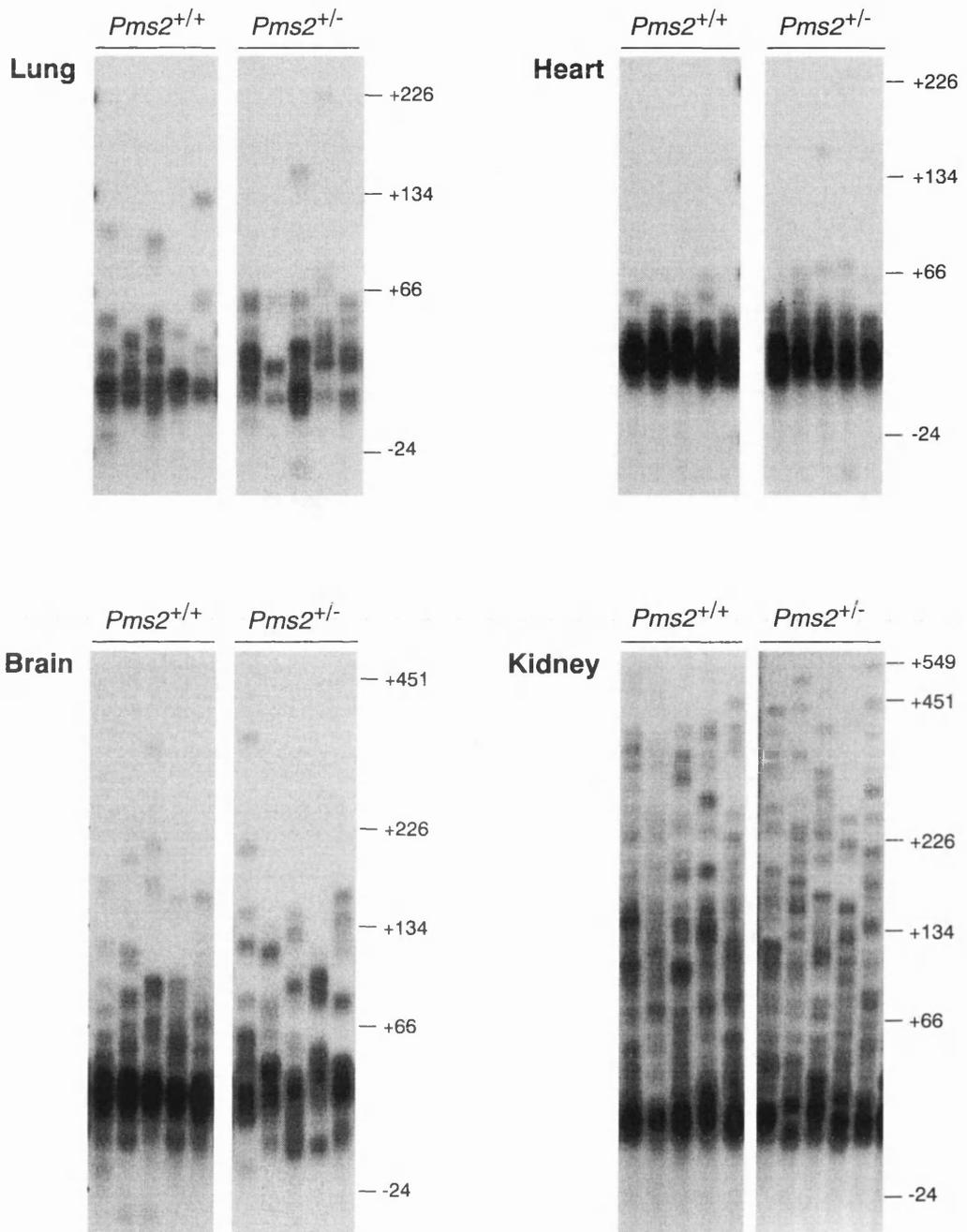
In order to assess the effect of the lack of one functional *Pms2* allele in the levels of somatic mosaicism detected in *Dmt-D* mice, SP-PCR analyses were performed on DNA samples collected from *Pms2*<sup>+/-</sup> mice and *Pms2*<sup>+/+</sup> age-matched controls. Four different tissues, which have shown different levels of trinucleotide somatic mosaicism (Fortune *et al.*, 2000), were selected for this study, in order to cover the entire range of trinucleotide repeat instability that a tissue can display. Lung and heart were chosen given the very low levels of repeat instability they exhibit, whereas the kidney was selected as the tissue where the transgene exhibits the most dramatic repeat length changes. Brain was also included in this study, since it displays moderate region-specific trinucleotide repeat instability, and also because it is one of the more commonly affected tissues in CAG•CTG expansion disorders.

Initially, two 24-month-old littermates were studied: one carrying both undisrupted *Pms2* alleles (*Pms2*<sup>+/+</sup>), and one heterozygous for the *Pms2* deletion (*Pms2*<sup>+/-</sup>). Preliminary amplification of 20-100 transgene molecules per reaction failed to detect any major difference in the levels of somatic mosaicism observed between the two mice, for any of the tissues studied (Figure 8.3). To quantify this observation, single molecule analysis was carried out (Section 2.7.5). Following accurate sizing of a large number of single molecules for each tissue (between 109 and 313 molecules), repeat size distributions were determined (Figure 8.4). Once again, no major differences in the mutation profiles were detected between the *Pms2*<sup>+/+</sup> and the *Pms2*<sup>+/-</sup> mice.

The percentage of alleles differing more than ten repeats in either direction from the progenitor repeat might be taken as an approximate estimate of the levels of repeat instability in a particular tissue: the higher the percentage, the more unstable the repeat. Around half of the alleles sizes fell within this range in the lung of both mice: 50% in the *Pms2*<sup>+/+</sup> mouse, and 55% in the *Pms2*<sup>+/-</sup>. On the other hand, 74% of the transgene molecules collected from the kidney differed more than 10 repeats in size from the progenitor allele for both genetic backgrounds. A two-tailed Mann-Whitney *U* test has indeed revealed that the median repeat length change from progenitor allele is not statistically different between these two animals, not only for lung ( $p=0.54$ ), but also for kidney ( $p=0.85$ ). The apparent, but not statistically significant, higher levels in repeat instability observed in the lung of a *Pms2*<sup>+/-</sup> animal, may have resulted from the longer progenitor allele size inherited by this mouse: 196 repeats for *Pms2*<sup>+/-</sup>, versus 176 repeats for *Pms2*<sup>+/+</sup> as determined by SP-PCR analysis of tail DNA at weaning.

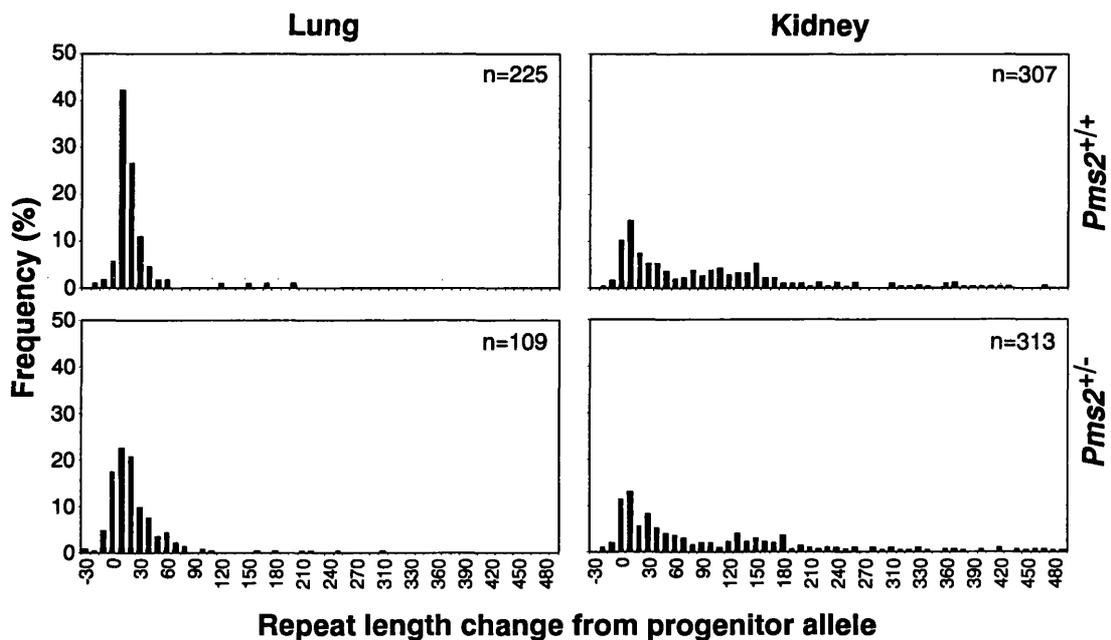
The observation that *Pms2*<sup>+/-</sup> mice show similar levels of somatic mosaicism to those detected in *Pms2*<sup>+/+</sup> animals, was further confirmed by the analysis of four other animals, aged 13 months, with progenitor allele sizes ranging from 150 to 169 repeats (Figure 8.5). High DNA input SP-PCR analyses corroborated that the two MMR genetic backgrounds failed to reveal major differences in the levels of somatic mosaicism detected in any of the four tissues analysed.

In summary, *Pms2*<sup>+/+</sup> and *Pms2*<sup>+/-</sup> mice exhibit very similar levels of trinucleotide repeat somatic mosaicism in all the tissues included in this analysis, as revealed by high DNA input SP-PCR, and by single molecule analysis. Indeed, these results confirm that the tissue-specific pattern



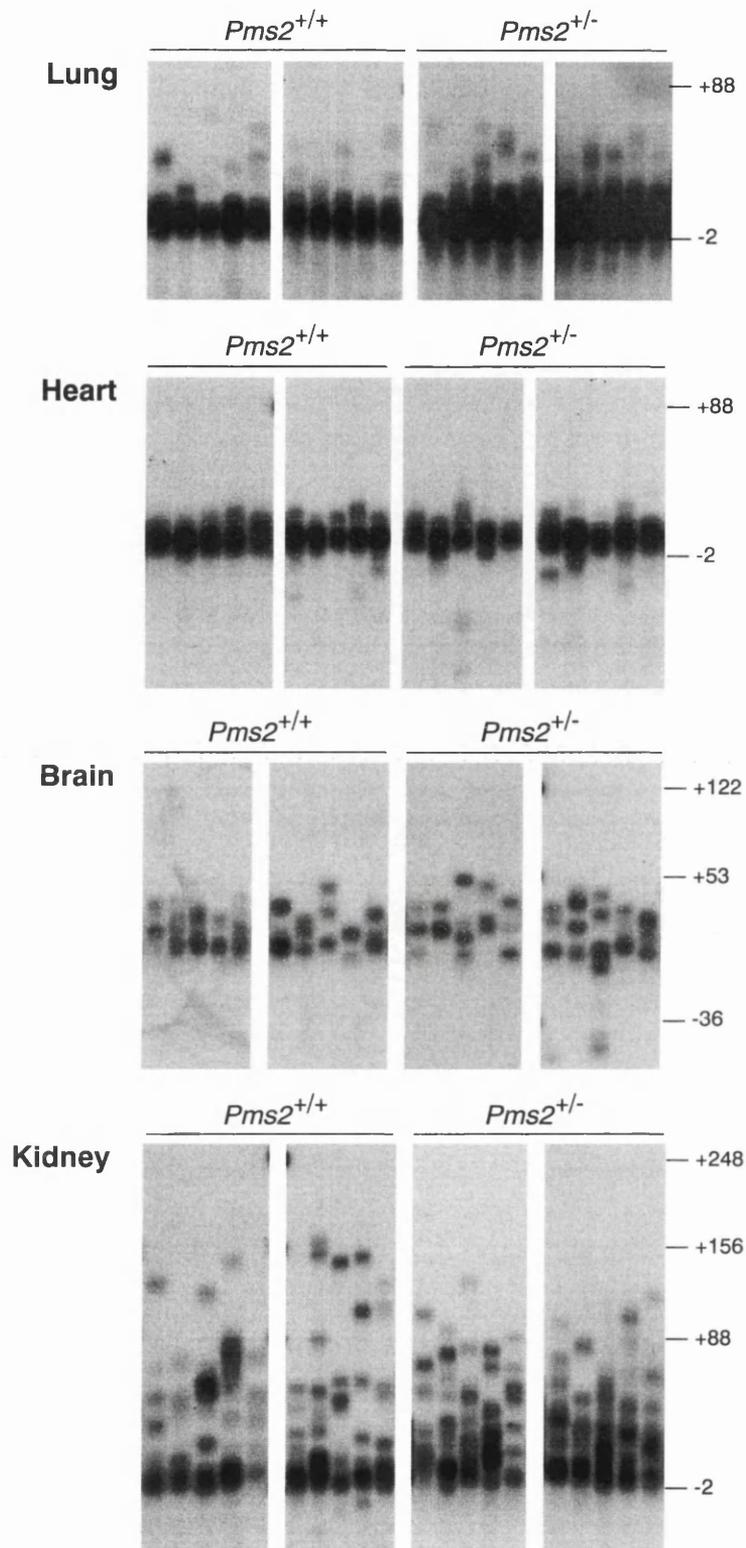
**Figure 8.3. Somatic mosaicism in 24-month-old *Dmt-D* mice heterozygous for the *Pms2* deficiency.**

Representative SP-PCR amplifications of DNA molecules extracted from lung, heart, brain and kidney, collected from two *Dmt-D* littermates aged 24 months. The mice differed on the number of functional *Pms2* allele copies: either one (*Pms2*<sup>+/-</sup>) or two (*Pms2*<sup>+/+</sup>). Around 20 transgene molecules were amplified from DNA samples collected from lung, whereas 50-100 molecules were amplified for all the other tissues shown. The size markers, converted into repeat size differences from the progenitor allele, are displayed on the right.



**Figure 8.4. Quantification of the effect of the disruption of one *Pms2* allele on the levels of trinucleotide somatic mosaicism detected in the lung and kidney of *Dmt-D* mice.**

Repeat size distributions in the lung and kidney of two *Dmt-D* littermates at 24 months of age. A mouse heterozygous for the *Pms2* disruption (*Pms2*<sup>+/-</sup>) was compared to its control littermate, which carried both copies of the same gene (*Pms2*<sup>+/+</sup>). SP-PCR analyses were performed with a very low DNA input: an average of 1-2 transgene molecules were amplified per reaction. Each band observed was individually sized and the repeat lengths grouped into 10 repeat size intervals. The total number of transgene molecules sized for each tissue (n) are shown in the top right corner of each panel.



**Figure 8.5. Somatic mosaicism in 13-month-old *Dmt-D* mice heterozygous for the *Pms2* deletion.**

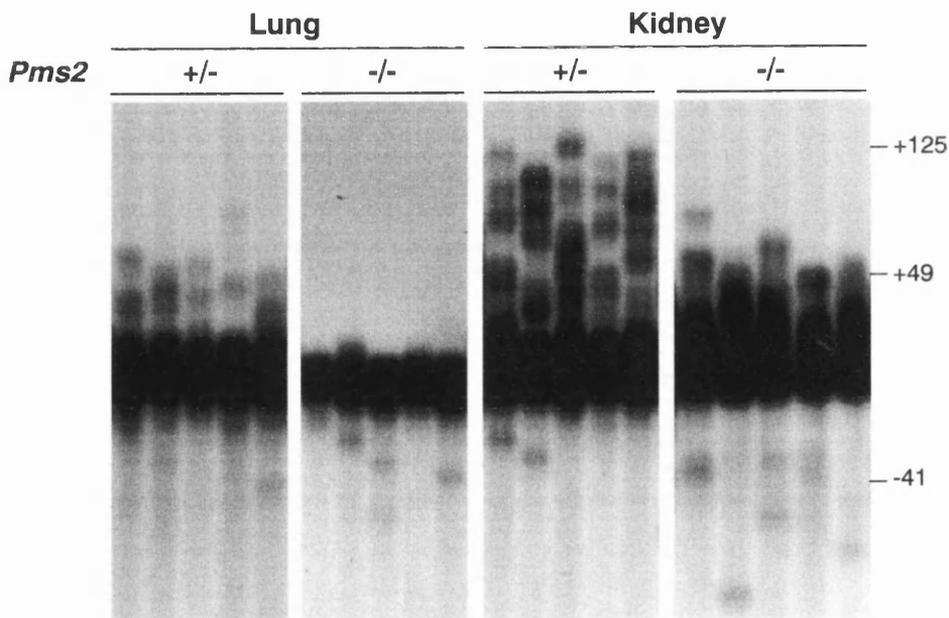
Representative SP-PCR analyses of transgene molecules extracted from lung, heart, brain and kidney of four *Dmt-D* mice aged 13 months. The two *Pms2*<sup>+/-</sup> mice carried a disrupted a *Pms2* allele, and the trinucleotide somatic instability they exhibited was compared with two age-matched controls, carrying two functional *Pms2* alleles (*Pms2*<sup>+/+</sup>). Around 10 transgene molecules from heart were amplified in each reaction, whereas 20-50 were amplified for the other three tissues. The size markers, converted into repeat number differences from the inherited allele are shown on the right.

of somatic mosaicism, usually observed in *Dmt-D* mice, is highly reproducible, and most certainly controlled by tissue- and/or cell-type *trans*-acting modifiers.

#### 8.2.4. Somatic mosaicism in homozygous *Pms2* knock-out mice

It should be mentioned that the assessment of somatic mosaicism in old *Pms2*<sup>-/-</sup> mice proved very difficult, since the total absence of the *Pms2* gene product in these animals made them ill, usually before six months of age. *Pms2*<sup>-/-</sup> mice developed paralysis of the hind legs, a condition that led to rapid deterioration and the subsequent sacrifice of these mice (Fortune, 2001).

The levels of somatic mosaicism in *Pms2*-nullizygous mice were initially assessed in an eight-month-old mouse, and compared to the levels of triplet repeat instability detected in its *Pms2*<sup>+/-</sup> age-matched control littermate. DNA samples from lung, heart, brain and kidney were analysed by SP-PCR. The amplification of ~50 transgene molecules per reaction revealed lower levels of somatic mosaicism in the lung and kidney of a homozygous *Pms2* knock-out mouse (Figure 8.6). The same result was observed for heart and brain samples (data not shown).<sup>7</sup>

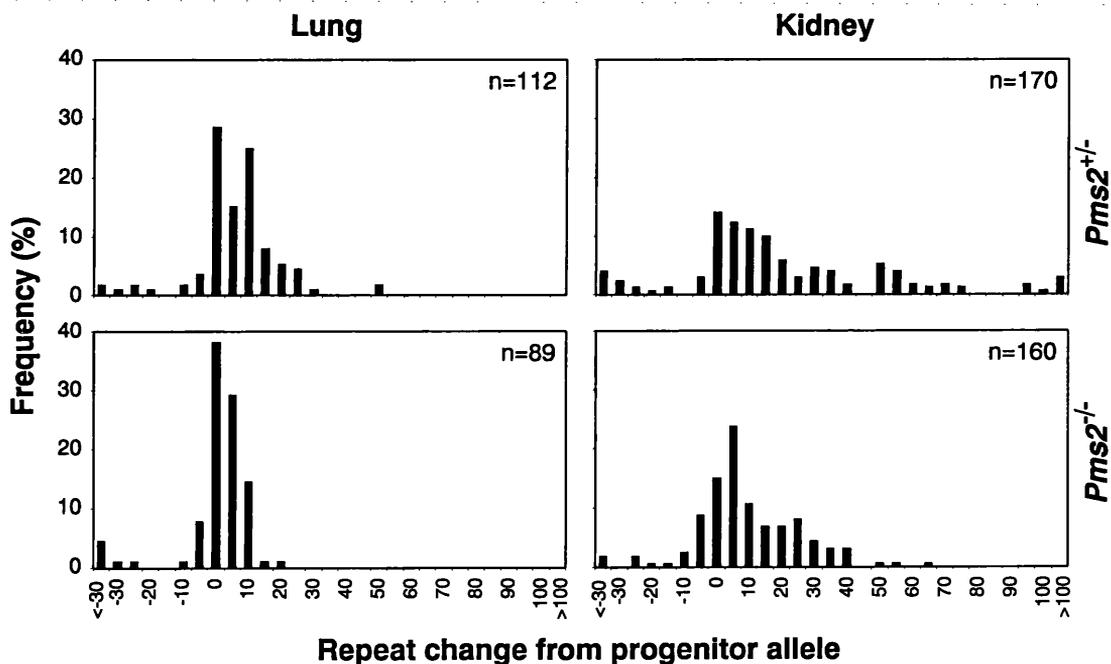


**Figure 8.6. Somatic mosaicism in lung and kidney of an eight-month-old *Dmt-D* mouse homozygous for the *Pms2* deletion.**

The autoradiographs shown in this figure represent SP-PCR analyses of transgene molecules extracted from lung and kidney tissue samples from two eight-month-old *Dmt-D* littermates: a heterozygous mouse for the *Pms2* disruption (*Pms2*<sup>+/-</sup>) and a homozygous for the same mutation (*Pms2*<sup>-/-</sup>). Around 50 transgene molecules were amplified in each reaction, using oligonucleotide primers DM-C and mDmtD-GR. The size markers were converted into repeat length differences from the progenitor alleles, and displayed on the right.

<sup>7</sup> High DNA input and single molecule SP-PCR analyses of DNA samples collected from the two eight-month-old *Dmt-D* mice, differing in their *Pms2* status, were performed by Laura Ingram, an honours undergraduate student at the University of Glasgow, under my supervision.

The decreased trinucleotide repeat instability in a *Pms2* knock-out genetic background was confirmed by single molecule analysis, performed on 89-170 transgene molecules collected from lung and kidney. The mutation profiles determined for these mice revealed lower levels of expansion-biased triplet repeat instability associated with the disruption of both *Pms2* alleles, in the two tissues analysed (Figure 8.7). While only 25% of the lung cells carried transgene sequences differing more than five repeats in size from the progenitor allele in the *Pms2*<sup>-/-</sup> mouse, 53% were included in the same repeat length range in its *Pms2*<sup>+/-</sup> littermate. The same comparison is also remarkable in the kidney of both animals: 52% in the *Pms2*<sup>-/-</sup> mouse, against 71% in the *Pms2*<sup>+/-</sup> animal (Figure 8.7). Furthermore, much greater repeat size ranges were detected in the tissues collected from the *Pms2*<sup>+/-</sup> mouse, relative to its *Pms2*-nullzygous littermate, which emphasises the effect of *Pms2* status on the repeat instability. Statistically, the median repeat length change from the progenitor in the heterozygous mouse was significantly larger than the median repeat length change in the homozygous littermate, for both lung ( $p=0.0007$ , two-tailed Mann-Whitney *U* test) and kidney ( $p=0.0038$ , two-tailed Mann-Whitney *U* test).



**Figure 8.7.** Quantification of the effect of a homozygous *Pms2*-deficient genetic background on the trinucleotide somatic mosaicism detected in the lung and kidney of *Dmt-D* mice.

Repeat size distributions in the lung and kidney of two *Dmt-D* littermates, aged eight months: a mouse heterozygous for the *Pms2* disruption (*Pms2*<sup>+/-</sup>) and a *Pms2*-null homozygote (*Pms2*<sup>-/-</sup>). SP-PCR analyses were performed with a very low DNA input: an average of 1-2 transgene molecules were amplified per reaction, using oligonucleotide primers DM-C and mDmtD-GR. Each band observed was individually sized and the repeat lengths grouped into 10 repeat size intervals. The total number of transgene molecules sized for each tissue (n) are shown in the top right corner of each panel.

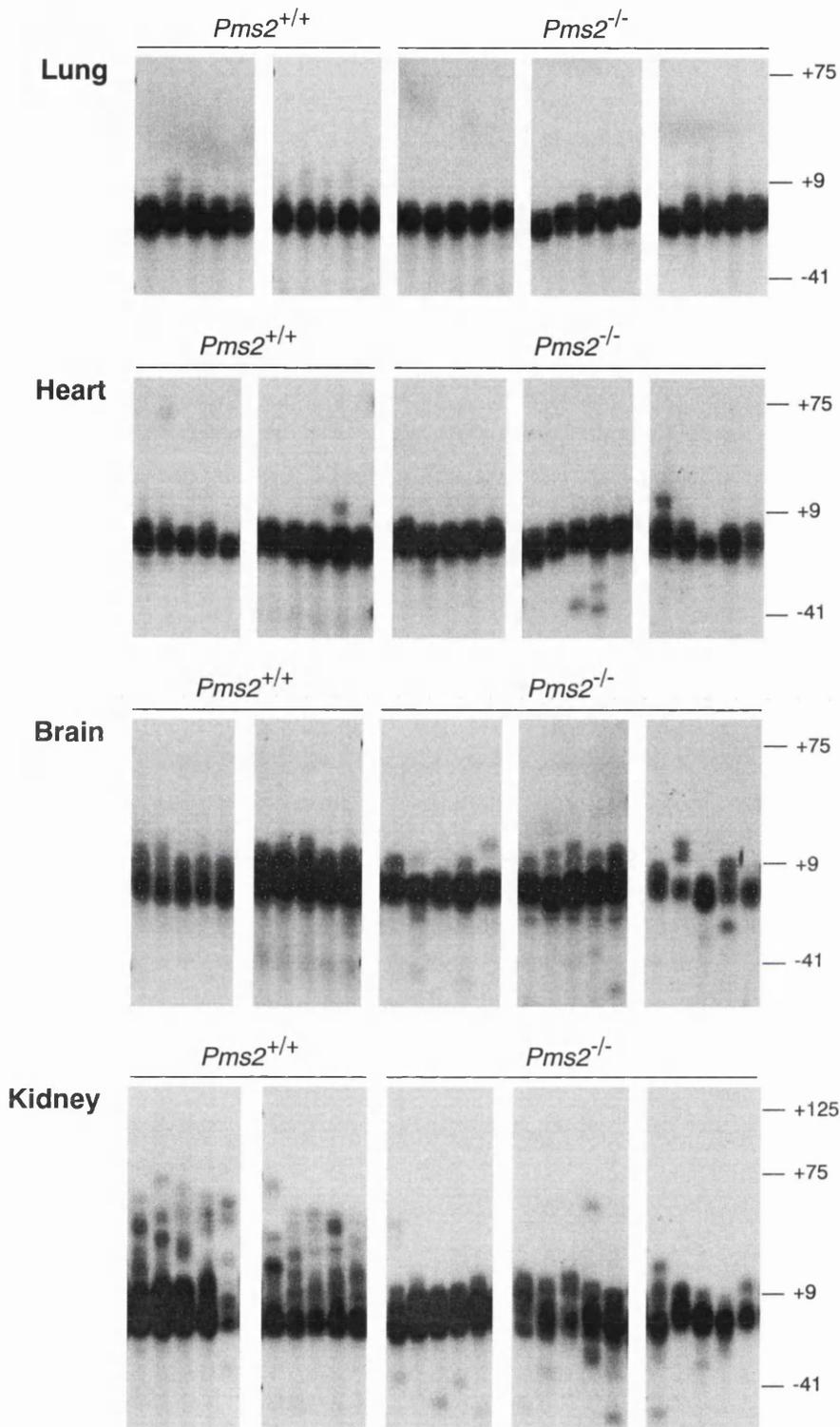
Since no statistically significant differences were found between *Pms2*<sup>+/+</sup> and *Pms2*<sup>+/-</sup> mice, the study of the effects of *Pms2*-deficiency on the dynamics of expanded trinucleotide repeat tracts was further extended to another five six-month-old animals, with inherited repeat lengths ranging from 166 to 176 CAG•CTG units. The levels of somatic instability in two *Pms2*<sup>+/+</sup> and three *Pms2*<sup>+/-</sup> animals were assessed by SP-PCR amplification of 20 to 50 transgene molecules collected from lung, heart, brain and kidney (Figure 8.8). Lower levels of somatic mosaicism were consistently detected in the kidney of *Pms2*<sup>+/-</sup> mice. A subtle but possibly significant difference may also be observed between lung DNA samples, with higher expansion rates being detected in *Pms2*<sup>+/+</sup> animals. However, this type of analysis was inconclusive for the two other tissues analysed (heart and brain). At six months of age trinucleotide somatic mosaicism in tissues other than kidney is hardly detected, and it is only then that tissue-specific repeat instability profiles begin to develop. Therefore, any qualitative observations in tissues other than kidney are very preliminary at this stage. Absolute conclusions could only be drawn from these experiments if quantification of the levels of repeat stability was carried out by single molecule SP-PCR analysis. Nevertheless, qualitative SP-PCR analysis of kidney transgene molecules was sufficient to conclude that *Pms2* nullizygous mice display lower levels of repeat instability.

To summarise, *Dmt-D* mice that lack both functional *Pms2* alleles exhibited lower levels of somatic repeat instability compared to their age-matched controls, carrying either one or two undisrupted *Pms2* copies.

### **8.2.5. Quantification of large repeat length mutations in a *Pms2*-deficient background.**

During the course of this study, it became evident that the amplification of high numbers of transgene molecules per SP-PCR (more than 100 template molecules), revealed marked accumulation of rare, but large (longer than 20 or 30 repeats) repeat length mutations (either expansions or contractions) in *Pms2*-null homozygotes. To confirm these findings, high DNA input SP-PCR analyses were performed on heart and brain DNA samples, collected from two eight-month-old littermates with different *Pms2* status, either heterozygous or homozygous for the *Pms2* gene disruption (Figure 8.9). A higher frequency of large repeat deletions was indeed observed in both tissues of the homozygous mouse for the *Pms2* mutation. Similar results were observed in the lung and kidney of the same mice (Figure 8.6). At six months of age, the same conclusion could be drawn from high DNA SP-PCR amplifications (Figure 8.8). However, at this earlier stage, only kidney DNA samples exhibited a detectable accumulation of large deletions in *Pms2*<sup>+/-</sup> mice, relative to their *Pms2*<sup>+/+</sup> age-matched controls (Figure 8.8).

To determine the difference in the frequency of large deletion events at eight months of age between *Pms2*<sup>+/-</sup> and *Pms2*<sup>+/-</sup> mice, 150-200 transgene molecules extracted from different tissues were amplified per reaction, in order to reveal the rare but large repeat length changes in lung, heart, brain and kidney. Following amplification of 3200-5000 transgene molecules, the frequency of these rare mutation events was calculated (Figure 8.10). The levels of large deletions showed a

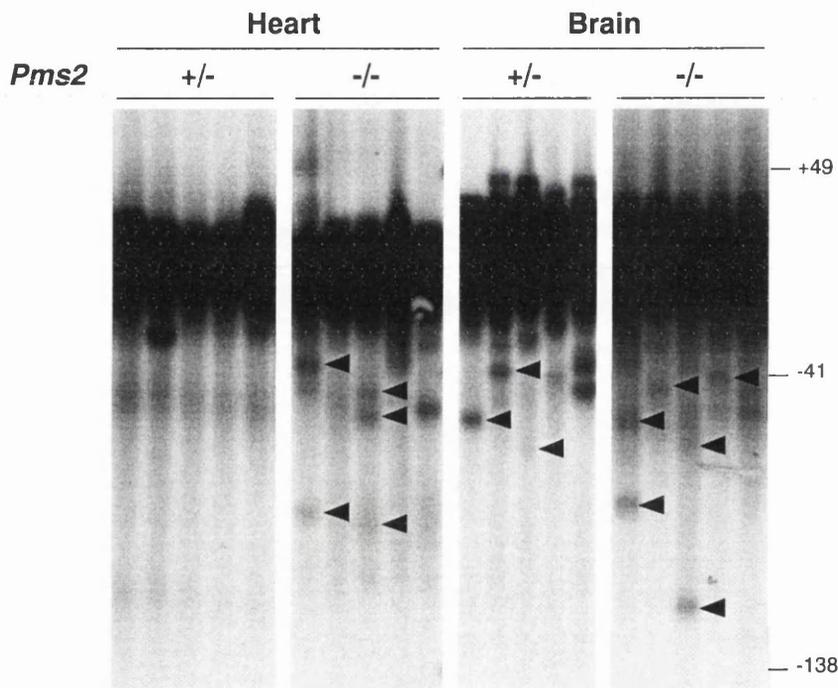


**Figure 8.8. Somatic mosaicism in six-month-old *Dmt-D* mice homozygous for the *Pms2* deletion.**

Representative SP-PCR analyses of transgene molecules extracted from lung, heart, brain and kidney harvested from five *Dmt-D* mice aged six months. Three *Pms2*<sup>-/-</sup> mice carrying a disruption in both *Pms2* alleles were analysed, and compared with two age controls, carrying two functional copies of the same gene. Twenty to 50 transgene molecules were amplified in each reaction. The size markers, converted into repeat size differences from the progenitor allele, are displayed on the right.

three-fold increase in the heart, and a two-fold increase in the lung and kidney of a *Pms2*<sup>-/-</sup> mouse, when compared to its *Pms2*<sup>+/-</sup> littermate. Moreover the shortest molecules amplified were around 10 to 30 repeats shorter in the tissues collected from the homozygous mouse, except for lung. The median large deletion in the brain of a *Pms2* nullizygous mouse was significantly lower than the median large deletion for the same tissue in a heterozygous littermate ( $p=0.0118$ , two-tailed Mann-Whitney *U* test), meaning that in the absence of *Pms2*, longer trinucleotide repeat contractions occur.<sup>8</sup>

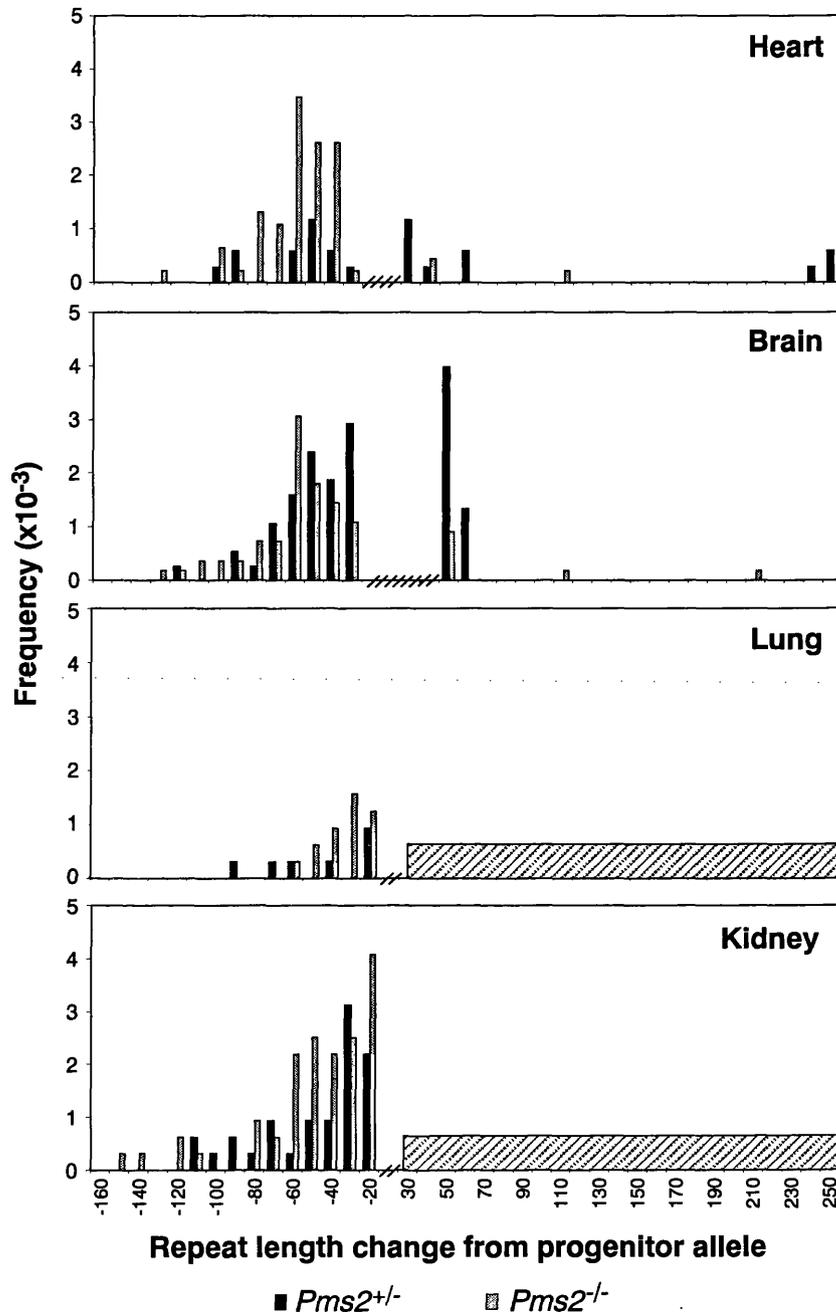
In conclusion, in the total absence of *Pms2*, the frequency of rare large deletions (>30 repeats) was increased. Nonetheless, these still only occurred in a small subset of cells (~1/1,000).



**Figure 8.9. High DNA input SP-PCR amplifications for the quantification of large repeat number changes in the heart and brain of a *Pms2*-null homozygous mouse.**

The amplification of high amounts of DNA in each reaction allowed the detection of large but rare repeat size changes (some examples pointed out in the figure by black arrowheads, ◄). Around 150 to 200 transgene molecules extracted from heart and brain of two eight-month-old *Dmt-D* mice, were amplified in multiple independent reactions with oligonucleotide primers DM-C and mDmtD-GR, giving rise to the SP-PCR products shown. The size markers, converted into repeat size differences from the progenitor allele, are displayed on the right.

<sup>8</sup> The quantification of large deletions in *Dmt-D* heart and brain tissue samples was also performed by Laura Ingram, under my supervision.



**Figure 8.10. Frequency of large trinucleotide repeat number mutants in the brain, heart, lung and kidney in a *Pms2*-deficient background.**

The frequency of large repeat size mutations was determined in two eight-month-old *Dmt-D* littermates: a mouse heterozygote for the *Pms2* deletion, *Pms2*<sup>+/-</sup> (■), and a homozygote for the same mutation, *Pms2*<sup>-/-</sup> (▨). A total number of 3200-5500 transgene molecules were amplified at high DNA concentration (100-200 molecules per reaction) with oligonucleotide primers DM-C and mDmtD-GR. All large deletions and expansions detected were individually sized for each tissue. Quantitative analysis of repeat expansions in lung and kidney tissue samples is shown in Figure 8.7. Allele lengths were grouped into 10 repeat size ranges.

### 8.3. Discussion

The dynamics of trinucleotide repeats in the mammalian context is an intriguing issue, which may depend on the complex interaction between multiple factors. Although some deductions can be drawn from studies carried out with simple model organisms, such as bacteria and yeast, it appears that the mammalian system sits on a higher level of complexity. One should therefore be aware of the differences and limitations of using simple organisms to model the mammalian system, since the conclusions and hypotheses inferred may not entirely apply to higher eukaryotes.

It has been suggested that the MMR system plays a central role in determining the stability of trinucleotide repeats in bacteria (Jaworski *et al.*, 1995; Parniewski *et al.*, 2000; Schmidt *et al.*, 2000; Schumacher *et al.*, 1998) and in yeast (Rolfmeier *et al.*, 2000; Schweitzer and Livingston, 1997). MMR enzymes correct non-native DNA structures that arise primarily during DNA replication. These aberrant structures include incorrectly paired bases as well as insertion/deletion loops, which may form as a result of microsatellite instability. We have decided to investigate the relationship between DNA repair activity and levels of somatic instability in a murine model of trinucleotide repeat instability, by breeding *Dmt-D* mice (Monckton *et al.*, 1997) onto a *Pms2*-deficient background (Baker *et al.*, 1995).

Given the impaired repair activity in cells heterozygous for MMR mutations (Edelmann *et al.*, 1996; Marra *et al.*, 2001), we first considered the hypothesis that *Pms2*<sup>+/-</sup> mice would show levels of somatic mosaicism different from those detected in wild-type animals. However, the results showed that *Pms2*<sup>+/+</sup> and *Pms2*<sup>+/-</sup> mice exhibited similar levels of repeat instability at 13 and 24 months of age. Single molecule SP-PCR analysis has confirmed that the repeat distributions at 24 months of age were not statistically different between the two genetic backgrounds. These results indicated that one *Pms2* allele is sufficient to generate levels of trinucleotide repeat instability in somatic cells that are indistinguishable from those that develop when both *Pms2* copies are functional.

Whereas a single *Pms2* allele is sufficient to maintain the same high levels of somatic mosaicism than those observed in mice carrying a totally functional MMR system, the disruption of both *Pms2* alleles causes a dramatic stabilisation of the triplet repeat tract, indicating that *Pms2* acts as a major genetic modifier of trinucleotide repeat instability in somatic tissues. All *Pms2*<sup>-/-</sup> mice analysed in this study displayed lower levels of somatic mosaicism than their *Pms2*<sup>+/+</sup> or *Pms2*<sup>+/-</sup> age-matched controls, in all the tissues analysed. Single molecule analysis revealed that the difference between *Pms2*<sup>+/-</sup> and *Pms2*<sup>-/-</sup> was statistically significant at eight months of age. This result was qualitatively confirmed by the comparison of somatic mosaicism between *Pms2*<sup>+/+</sup> and *Pms2*<sup>-/-</sup> mice at six months of age. In all cases higher levels of repeat instability were observed in the presence of a fully functional MMR system, in particular for lung and kidney.

Taken together, these observations clearly suggest that *Pms2* may be considered a genetic enhancer of trinucleotide repeat instability in the soma. This situation is in clear contrast with most studies, previously carried out in simple model organisms. In bacteria, *mutS* and *mutL* mutations cause an overall increase in the frequency of small length changes within trinucleotide repeat tracts

(Parniewski *et al.*, 2000; Schmidt *et al.*, 2000; Schweitzer and Livingston, 1997). However, an increased frequency of large deletions within long trinucleotide repeat tracts cloned into MMR-proficient *E. coli* strains has been reported (Jaworski *et al.*, 1995; Parniewski *et al.*, 2000; Schmidt *et al.*, 2000). In yeast, both *pms1* and *msh2* mutations destabilise trinucleotide repetitive arrays (Schweitzer and Livingston, 1997), *pms1* being the yeast homologue of mammalian *PMS2* (Nicolaides *et al.*, 1994). It appears that trinucleotide repeat dynamics is controlled by different mechanisms in different organisms, or that the same modifiers (such as MMR proteins) may play different roles in different contexts.

If replication slippage were the primary mechanism by which somatic mosaicism arose, then it would be expected that MMR loss of function would enhance repeat instability. In this particular case, if *Pms2* monitored and prevented trinucleotide repeat expansions, the *Pms2* deletion would automatically impair the repair of slippage errors, thereby causing increased levels of somatic mosaicism. However, the data presented in this chapter strongly argues against such a mechanism. It is therefore conceivable that trinucleotide repeat expansion relies on recognition and binding of MMR proteins to alternative DNA structures (Pearson *et al.*, 1997) formed within the repeat sequences (Pearson and Sinden, 1998b), followed by the inappropriate repair (or “misrepair”) of hairpins and loop-outs. Rather than being considered as lesions that have escaped MMR, trinucleotide repeat expansions may be viewed as the outcome of “misrepair” by at least some components of the MMR system. Alternatively, it is theoretically possible that the greater triplet repeat stability observed in *Pms2*<sup>-/-</sup> mice might be explained by preferential repair of newly synthesised DNA, as opposed to the template strand, mediated by an underlying or competitive repair pathway still active in the absence of *Pms2*.

However, one important point should be made. Despite being highly sensitive, and able to detect subtle changes in repeat number, SP-PCR techniques cannot identify the origin of such mutations. It is therefore conceivable, that *Pms2*<sup>-/-</sup> mice, being defective in MMR activity, exhibit a greater accumulation of small mutations due to DNA polymerase slippage, which cannot be properly repaired in an MMR-deficient genetic background. These mutations are most likely to involve only a very few repeat units and cannot be easily identified by SP-PCR procedures. In contrast with an overall mutation rate, the actual frequency of replication slippage errors cannot be determined. Indeed, increased instability of short CAG•CTG repeats at both *DMI* and *SBMA* loci have been described in tumour samples from HNPCC families (Wooster *et al.*, 1994). Similarly, the mutation rates of expanded *DMI* alleles in ovarian and gastrointestinal tumour tissue samples were higher than in normal tissue samples from the same organ (Jinnai *et al.*, 1999; Kinoshita *et al.*, 1997). This is in agreement with the models proposed, in order to resolve the opposing effects of MMR mutations on the dynamics of CTG•CAG repeat tracts cloned into *E. coli* (Parniewski *et al.*, 2000; Schmidt *et al.*, 2000). Small contractions and deletions, within short trinucleotide repeat sequences are generated by DNA polymerase slippage and unrepaired in the absence of MMR, leading to the accumulation of small repeat changes. In contrast, following a strand slippage event within a long trinucleotide repeat tract, the excision of the daughter strand by a functional MMR system, facilitates single-strand folding into secondary structures within the template DNA

sequence, thereby promoting large deletions (Jaworski *et al.*, 1995; Schmidt *et al.*, 2000). By analogy, it is certainly likely that in the absence of Pms2, strand slippage-mediated mutations may accumulate in dividing cells, but they are certainly not key determinants in the development of the dramatic expansion-biased repeat profiles observed in tissues from *Dmt-D* mice. However, contradicting the effect of MMR system in promoting large deletions in *E. coli* (Jaworski *et al.*, 1995; Parniewski *et al.*, 2000; Schmidt *et al.*, 2000) in the total absence of Pms2, a greater frequency, yet still very low, of large deletions was detected in *Dmt-D* mice, implying a role for *Pms2* in avoiding large repeat number changes within trinucleotide sequences. Similarly, the correction of very large looped mispairs in yeast involves, at least partially, *msh2* and *pms1* gene products (Clikeman *et al.*, 2001), and it has also been reported that *pms1*-deficient yeast also accumulate long contractions (Schweitzer and Livingston, 1997). It may be hypothesised, that in the absence of *Pms2*, unusually long deletions may arise through the folding back of the CAG•CTG repeat tract. Under MMR-deficient conditions, single-stranded DNA secondary structures would be alternatively processed, leading to a large repeat contraction, possibly through a bypass of the structure during repair DNA synthesis.

In summary, MMR proteins may actually be active participants in the mutation mechanism that drives triplet repeat instability. Interestingly, *Pms2* has been shown to be necessary for the accumulation of highly mutated sequences in the variable region of immunoglobulin genes (Kim *et al.*, 1999), indicating that trinucleotide repeat diseases are not the first situation in which fully functional MMR components are involved in a mutational process.

*Msh2* has also been reported as being required for the development of somatic instability of an expanded CAG•CTG repeat tract in a transgenic mouse model of Huntington disease, carrying an expanded exon 1 of the human *HD* gene (Manley *et al.*, 1999b). Similar findings have been described for *Msh3* and somatic instability levels of an expanded CTG•CAG repeat inserted into the murine *Dmpk* gene by homologous recombination (van Den Broek *et al.*, 2002). The effects of *Msh2* and *Msh3* mutation observed in those mice appear to be more dramatic than the effect of *Pms2*-disruption reported here. In both studies electrophoretic profiles generated by bulk DNA amplification, using fluorescent primers, automated fragment analysis and GeneScan software, were compared between MMR-proficient and MMR-deficient mice. It is now well established that GeneScan-based techniques are not as sensitive as SP-PCR analysis, leading to an underestimation of the levels of repeat instability (Fortune *et al.*, 2000; Kennedy and Shelbourne, 2000). This may account for the apparent greater effects of *Msh2* and *Msh3* mutations on the CAG•CTG tract dynamics.

Nevertheless, the specific role of individual MMR genes in the repair pathway may have a greater impact on the different extents to which individual MMR gene mutations affect trinucleotide repeat instability. Indeed, in contrast to the effects of *Msh2*, *Msh3* and *Pms2* mutations, a functional *Msh6* gene has been found to stabilise trinucleotide repeat tracts in the somatic cells of DM1 knock-in mice (van Den Broek *et al.*, 2002), strongly suggesting that different MutS homologues have distinct functions in trinucleotide repeat metabolism (for further details see Chapter 9). Mononucleotide sequences have shown higher mutation frequencies in mice

lacking *Msh2*, than in those deficient in *Pms2* (Andrew *et al.*, 2000). This is interesting in view of the fact that *Pms2*<sup>-/-</sup> mice are predisposed to cancer, mainly lymphomas, but do not develop gastrointestinal neoplasias (Baker *et al.*, 1995; Prolla *et al.*, 1998), whereas *Msh2*<sup>-/-</sup> and *Msh3*<sup>-/-</sup> animals do (Edelmann *et al.*, 2000; Reitmair *et al.*, 1995). The differences with respect to tumour phenotype between *Msh2*<sup>-/-</sup>, *Msh3*<sup>-/-</sup> and *Pms2*-deficient mice may be indicative of differential roles for Msh2, Msh3 and Pms2 proteins in mismatch identification and error correction, and account for the lower incidence of *PMS2* germline mutations in HPNCC patients (Peltomaki, 2001b). While MSH2 is involved in the repair of single base mismatches and insertion-deletion loops, MSH3 is specifically involved in the correction of insertion-deletion loops. In the absence of MSH2 or MSH3, loop-outs are very inefficiently corrected (Jiricny, 2000; Peltomaki, 2001a). After the characterisation of *MLH3*, some degree of redundancy between *PMS1*, *PMS2* and *MLH3* has been suggested (Lipkin *et al.*, 2000). Both MLH1-MLH3 and MLH1-PMS2 heterodimers can specifically repair insertion-deletion loops. Therefore, in the absence of PMS2, insertion-deletion loops can still be repaired by MLH1-MLH3 complexes. These findings may also help to explain why a mutation in mouse *Pms2* may have a weaker effect on genomic stability and cancer predisposition, compared to mutations in *Msh2* and *Msh3*, and why there are still detectable levels of trinucleotide repeat somatic mosaicism in the absence of *Pms2*.

Differences in the genetic background between our mice and the animals used in other studies may also explain the greater effect of *Msh2* and *Msh3* deletions, in the same way that *Msh6* deficiency may or may not induce intestinal tumours, depending on the mouse genetic background (de Wind *et al.*, 1999; Edelmann *et al.*, 1997), indicating a role for genetic or environmental modifiers in determining the outcome of an MMR mutation. The genetic background also appears to contribute to the levels of somatic mosaicism detected in a knock-in mouse model of DM1 (van Den Broek *et al.*, 2002).

Increasing evidence has suggested that the expansion mechanism of trinucleotide sequences is not entirely dependent on the pausing of DNA polymerase, followed by slippage and lack of repair of alternative structures, as the replication slippage model predicts (Richards and Sutherland, 1994). An alternative mechanism for trinucleotide repeat expansion has been proposed, suggesting that germline trinucleotide repeat expansion is controlled by repair-dependent replication machinery, rather than mitotic or meiotic events (Kovtun and McMurray, 2001). If instability were limited to mitotic populations, Msh2-dependent recognition and stabilisation of slipped-stranded DNA structures could lead to expansion following a subsequent round of replication. However, if the trinucleotide repeat mutation pathway were primarily dependent on MMR-mediated stabilisation of alternative DNA secondary structures, the changes in repeat number would still require cell division and DNA replication to occur. The high proportions of non-mitotic cells in brain tissues that show greatest instability in several mouse models (Fortune *et al.*, 2000; Kennedy and Shelbourne, 2000; Lia *et al.*, 1998; Mangiarini *et al.*, 1997; Seznec *et al.*, 2000) indicate that repair or some other non-mitotic DNA synthesis may be a mechanism of expansion. Therefore, several factors appear to play determinant roles in the stabilisation of triplet repeat sequences within the mammalian genome. The work presented here suggests that at least

*Pms2*, but quite possibly other components of the MMR system, act as genetic enhancers of somatic trinucleotide repeat instability. The involvement of the *Pms2* gene product in the accumulation of repeat length variation is of major significance, since it implies that not only MMR proteins that recognise and directly bind to mismatches (MutS homologues), but also MutL homologues are involved in the mutation mechanism. Therefore, a more general involvement of MMR enzymes in the dynamics of expanded triplet repeats, further than the simple stabilisation of loop-outs or hairpin-like structures, must be considered (a more detailed discussion on this topic and a possible molecular mechanism of trinucleotide repeat expansion are included in Chapter 9).

It is noteworthy, that the ability of human mismatch repair proteins to interact with other members of the rodent mismatch repair machinery indicates a high degree of conservation (Marra *et al.*, 1998), therefore it is not unreasonable to speculate that the findings reported here may also apply to humans, and that the MMR system may act as an enhancer of somatic mosaicism in trinucleotide repeat disorders.

## 9. Main conclusions, final discussion and future perspectives

DNA has long been known as the reservoir of genetic information, thought to be essentially passive and stable through numerous cellular processes. However, this picture has been changing. Since 1991, a fast-growing group of neurological and neuromuscular diseases have been associated with chain length expansions of trinucleotide repeats CAG•CTG, CGG•CCG, and GAA•TTC. Multiple short trinucleotide sequences are normally present in the genome. They are usually polymorphic among the normal population, and relatively stable in both germline and somatic cells. However, under certain circumstances, the copy number increases beyond a certain level, at which point the triplet repeat expansion becomes highly unstable and pathogenic, causing severe human conditions. The length of these repeats is closely related to the onset and progression of clinical symptoms. Disease-associated unstable trinucleotide repeats have been found in various regions of the affected genes, such as 3' and 5'-untranslated regions, exons and introns (Cummings and Zoghbi, 2000a; Cummings and Zoghbi, 2000b; Richards, 2001).

Although substantial progress had been made in understanding this fascinating mutational mechanism, new information continues to raise important questions. While a thorough understanding of the mechanism has not been achieved, the search is well worth the effort. A mechanistic solution to the problem of instability is likely to expose the very nature of heritable traits and provide valuable clues towards the development of promising therapeutical routes.

### 9.1. Main conclusions

#### 9.1.1. Trinucleotide repeat dynamics and cell division

In order to better understand the mutational pathway of trinucleotide repeat sequences in somatic cells, repeat length variation was assessed in cell cultures established from *Dmt* transgenic mice carrying CAG•CTG repetitive tracts in their genome (Monckton *et al.*, 1997). The site of integration of the transgene has proved to be a key determinant of trinucleotide instability, not only in the soma (Fortune *et al.*, 2000), but also in germline transmissions (Zhang *et al.*, 2002). Age-dependent, tissue-specific, expansion-biased somatic mosaicism had been previously detected in a particular transgenic line, the *Dmt-D* mice, carrying a single copy of the *Dmt* transgene (Fortune *et al.*, 2000).

SP-PCR techniques were used during the course of this project to analyse in detail the variation in the CAG•CTG repeat length, not only in *Dmt* mouse cell lines but also in mouse tissue samples. SP-PCR is an extremely sensitive technique capable of revealing subtle, but important differences in repeat number distributions, not detectable using bulk DNA analysis. SP-PCR

methods represent a robust approach for accurately assessing the level of repeat variation, providing both quantitative and qualitative data (Monckton *et al.*, 1995).

Monitoring the dynamics of expanded CAG•CTG repeats in *Dmt-D* cultured cells over long time periods, and numerous population doublings, revealed that these cultures were able to recreate expansion-biased trinucleotide repeat instability. In contrast to an apparent “synchronised” expansion, observed in EBV-transformed human lymphoblastoid cell lines (Khajavi *et al.*, 2001), *Dmt-D* cell lines have usually shown a continuous increase in the mean repeat size, similar to that detected in human tissue samples. Most importantly, the tissue-specific patterns of somatic mosaicism detected *in vivo* (Fortune *et al.*, 2000), were also retained in culture, suggesting that tissue- and cell-specific factors must play determinant roles in the process of somatic mosaicism. Intriguingly, cells carrying longer alleles appeared to preferentially accumulate in culture during the first few passages, possibly due to a selective growth advantage of cells that already contained longer repeats *in vivo* (Chapter 3).

EBV-transformed DM1 lymphoblastoid cell lines (LBCLs) with larger CTG•CAG repeat expansion appear to have a growth advantage over those with smaller expansions in culture, a phenomenon termed “mitotic drive” (Khajavi *et al.*, 2001). The authors proposed that the expansion of the DM1 CTG•CAG repeat downregulates p21, leading to increased LBCL cell proliferation and possibly to increased levels of apoptosis, and resulting in shortened cellular life spans (Khajavi *et al.*, 2001). p21 down-regulation could result directly from altered expression levels of one of the genes in the *DM1* locus (Khajavi *et al.*, 2001), or could instead be mediated by the recruitment of CUG-BP into the nucleus (Timchenko *et al.*, 2001b). The fact that cell proliferation rates of *Dmt-D* cell lines did not correlate with CAG•CTG repeat lengths, suggests that the intriguing association reported by Khajavi *et al.* cannot be uniquely accounted for by the presence an expanded CAG•CTG sequence at the DNA level. Since the *Dmt* transgene does not include the *DMPK* coding region, the effect observed in human DM1 LBCKs, could indeed result from the altered expression levels of one of the genes that map within the *DM1* locus. Alternatively, the parallel between repeat size and cell proliferation rate might be mediated by expanded CUG repeats within *DMPK* transcripts, again rendering it unlikely to be detected in our cell culture system. Reverse transcriptase PCR analysis revealed that the *Dmt162* transgene is ubiquitously transcribed into heterogeneous nuclear RNA in *Dmt-D* mice (even in those tissues where the repeat is stable), but is not present in mRNA transcripts, suggesting that it has probably integrated into a mouse intron. However, the CAG strand, as opposed to the presumed CUG toxic repeats, is preferentially transcribed (Fortune, 2001). In addition, no CUG-containing nuclear foci were detected in D2763L lung cells (J. Houseley and D.G. Monckton, personal communication).

Together with a basic expansion-biased mutation mechanism detected in cultured *Dmt-D* kidney cells (Chapter 3), an association between faster cell proliferation and larger CAG•CTG repeat expansions in DM1 cells (Khajavi *et al.*, 2001) is particularly attractive since it could contribute to the bias of CAG•CTG repeat towards further expansion in rapidly proliferating cells *in vivo*, but may not be relevant to mainly post-mitotic tissues, such as the muscle or the brain (Anvret *et al.*, 1993; Ashizawa *et al.*, 1993; Fortune *et al.*, 2000; Kennedy and Shelbourne, 2000;

Lia *et al.*, 1998; Mangiarini *et al.*, 1997; Monckton *et al.*, 1995; Seznec *et al.*, 2000; Telenius *et al.*, 1994; Thornton *et al.*, 1994; van Den Broek *et al.*, 2002).

*Dmt-D* fibroblastic cell lines may not represent a pathologically relevant cell type, but the generation of an *in vitro* system to study the complex pathophysiological cascade underlying the disease mechanism was never the purpose of this project. Although the continuous proliferation of mouse fibroblasts in this model system differs from the post-mitotic tissues, where the greatest expansions are usually detected *in vivo*, these cells still mimic stages of rapid proliferation, possibly similar to those occurring during embryogenesis, stem cell proliferation, or in pre-meiotic spermatogenic events, which may be important stages for the generation of significant repeat length variation (Leefflang *et al.*, 1999). Nevertheless, no correlation was found between the rates of cell turnover in homogenous *Dmt-D* cell lines, and the average repeat length or the degree of repeat size heterogeneity in culture (Chapter 3), possibly suggesting that the accumulation of allele length variability detected in cell lines undergoing rapid proliferation does not occur during DNA replication. In addition, accumulation of repeat length variation was still detected in arrested cells, sometimes at higher rates than in control cells proliferating under standard conditions (Chapter 7). In DM1 patients, the rates of expansion are indeed greater in slowly proliferative muscle cells than in more rapidly dividing leukocyte populations (Anvret *et al.*, 1993; Ashizawa *et al.*, 1993; Monckton *et al.*, 1995; Thornton *et al.*, 1994). Similarly, striking repeat expansions were reported in non-replicating striatal cells in *Dmt-D* mouse brains (Fortune *et al.*, 2000) and in a HD knock-in mouse model (Kennedy and Shelbourne, 2000). In addition, several genotoxic agents, which did not necessarily interfere with the rates of cell division, assessed by the population doubling time, were capable of interfering with the dynamics of expanded CAG•CTG repeats in culture (Chapters 4 and 6).

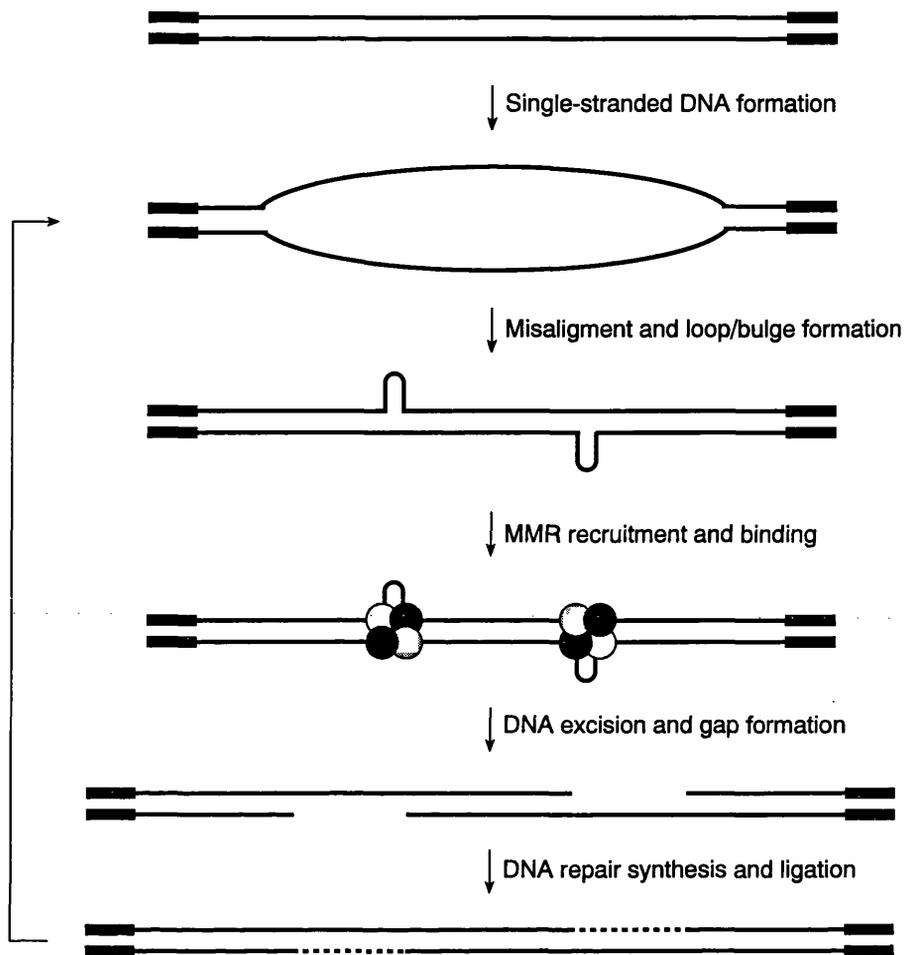
Combined, these data argue against a direct and simple link between cell division and the mutation process. The results described and discussed here do not preclude such a link, but indicate that other factors must be of critical importance. Indeed, the data presented in previous chapters does not exclude replication slippage as a possible mechanism for the generation of trinucleotide repeat length variability. The enhanced trinucleotide repeat dynamics in caffeine-treated cells, forced to enter mitosis through cell cycle checkpoint override, is consistent with the hypothesis that replication errors, generated by DNA polymerase slippage, can in fact mediate repeat expansion if not properly repaired (Chapter 6). The reduced rates of expansion in *Dmt-D* cells exposed to araC (Chapter 6) are also consistent with a role for DNA polymerase slippage in the expansion of trinucleotide repeat tracts. However, the decreased levels of somatic mosaicism in *Msh2* (Manley *et al.*, 1999b), *Msh3* (van Den Broek *et al.*, 2002) and *Pms2* null homozygous mice (Chapter 8) strongly suggest that replication slippage is not the major factor driving trinucleotide repeat instability in the soma. Moreover, replication slippage during genome duplication cannot certainly account for the accumulation of repeat length variation in non-dividing cells (Chapter 7). It certainly appears as though there are multiple components of the expansion-biased mutation process operating on trinucleotide repeat sequences, which may function in concert in most of the cells.

### 9.1.2. A possible model of trinucleotide repeat mutation

Few studies have addressed the molecular mechanisms underlying the dynamics of expanded trinucleotide repeats in post-mitotic cells. Messer and co-workers have reported that *Msh2* deletion dramatically decreases the levels of somatic mosaicism in the striatum of transgenic mice for an expanded exon 1 of the human *HD* gene (Manley *et al.*, 1999b). Moreover, *Msh2* is also required for repeat instability in developing sperm of the same mice (Kovtun and McMurray, 2001). These findings definitely implied MMR enzymes in the mechanism of trinucleotide repeat expansion and questioned the replication slippage mechanism. If replication slippage were the primary mechanism by which mosaicism arose, it would be expected that loss of function of MMR proteins would enhance repeat instability, since one of the primary functions of the mismatch repair machinery is to ensure the faithful replication of DNA during cell division.

Based on the ability of human MSH2 protein to bind slipped-stranded DNA structures (Pearson *et al.*, 1997), it was hypothesised that stabilisation of alternative secondary structures formed within triplet repeat tracts by MSH2 recognition and binding would lead to expansion following DNA repair synthesis (Section 1.4.5) (Kovtun and McMurray, 2001; Manley *et al.*, 1999b). However, it should be noted that following DNA gap repair, loops and slipped-stranded-like structures would be trapped in a heteroduplex DNA molecule (Kovtun and McMurray, 2001), and a second round of DNA replication and cell division would be required to complete the mutation process. Increasing experimental evidence strongly hints that repeat expansion can occur in the absence of cell division, as discussed in Chapter 7. Therefore, attractive as it is, the mechanism proposed by Kovtun and McMurray may only account for somatic trinucleotide repeat mutation in proliferating tissues, and is unlikely to contribute to the expansion of trinucleotide repeat sequences in post-mitotic cells.

In addition to a functional *Msh2* gene (Manley *et al.*, 1999b), both *Msh3* (van Den Broek *et al.*, 2002) and *Pms2* genes (Chapter 8) are also required for the accumulation of repeat size mosaicism of expanded CAG•CTG repeats in different transgenic mouse models. The finding that the *Pms2* gene product is also involved in the mutation mechanism is of major significance, as it implies that not only MMR proteins that directly recognise and bind to DNA mismatches are responsible for changes in the repeat number. The involvement of *Pms2* argues against the stabilisation of loop-outs and hairpins as the driving force of trinucleotide repeat mutation. Consequently, a mutational process relying on error-prone DNA repair, rather than on the simple binding of MMR components to secondary structures, presents a valid alternative hypothesis (Figure 9.1). The substrate for repair could either be damaged DNA, induced by reactive oxygen species (ROS) or other exogenous genotoxic agents, or non-orthodox DNA conformations. Alternative secondary structures, likely to form following dissociation of the double-stranded DNA duplex (not necessarily during DNA replication), might recruit functional components of the MMR machinery. In the absence of a strand discrimination signal, the MMR enzymes might mediate an expansion-biased repair pathway, by incorporating a newly synthesised DNA sequence opposite the loop, hairpin or DNA bulge.



**Figure 9.1. Mechanistic model of somatic trinucleotide repeat expansion mediated by MMR proteins.**

A double-stranded trinucleotide DNA sequence unpairs to generate two single-stranded DNA trinucleotide sequences (thin lines). The unpaired region may or may not extend to the flanking non-repetitive DNA sequences (thick lines). Following denaturation, single-stranded trinucleotide repeats may fold into alternative secondary structures, such as loop-outs and hairpins, stabilised by intrastrand base pairs. Alternatively, DNA strands may fail to align correctly, giving rise to small bulges or loop-outs. The MMR proteins recognise and bind alternative secondary structures, and in the absence of a strand-discrimination signal, may mediate the incorporation of newly synthesised DNA sequences (dashed lines) in the opposite DNA strand, leading to expansion following DNA ligation. The process may be repeated if the DNA duplex undergoes a second round of denaturation. The mechanism is not cell division-based, depending instead on the tendency for double-stranded DNA denaturation, the stability of putative secondary structures and activity of MMR enzymes.

This mutation mechanism may not only operate in proliferating cells during cell cycle stages other than the S phase, but also in non-dividing cells, as long as DNA damage or structural changes target MMR components to the repeat tract, and elicit a repair event. Although there is no evidence to support it, trinucleotide repeat tracts, or the alternative structures they form, may be particularly sensitive to damage, acting as a mutation hotspot, and therefore particularly dependent on DNA repair. Indeed, brain ageing is associated with the accumulation of DNA damage, thus DNA repair events may be increasingly important with age in neurons (Mrak *et al.*, 1997), accounting for the time-dependent accumulation of larger expansion in the brain of transgenic and knock-in HD mice (Kennedy and Shelbourne, 2000; Mangiarini *et al.*, 1997). Alternatively, a MMR response may be initiated by the simple formation of alternative structures following double strand DNA denaturation. Metabolic processes such as transcription, would create conditions to the formation of simple bulges, loop-outs or hairpins, which might be readily recognised by MMR enzymes, and engaged in a DNA repair pathway. Nevertheless, an association between somatic mosaicism and *DMPK* transcriptional activity has not been found (Lia *et al.*, 1998). In contrast, transgenic mice for the first exon of the *HD* gene have shown an association between transgene expression and relative stability, as the most stable line does not express the transgene (Mangiarini *et al.*, 1997). However, this could reflect a requirement for an open chromatin structure, rather than for transcription *per se* (Bates and Davies, 1997). Non-repetitive sequences immediately flanking the trinucleotide sequences may not only influence the propensity to form alternative structures, but may also have an effect on their stability. A high GC content, associated with higher mutation rates in humans (Brock *et al.*, 1999), might facilitate the folding of trinucleotide sequences into alternative DNA conformations, by undergoing faster reassociation, leaving two single-stranded trinucleotide sequences free to form intrastrand base pair and generate loop-outs or hairpins. In addition, once formed, alternative structures may also be more easily stabilised by a high GC content, given the higher melting temperature of GC-rich DNA sequences. Finally, it might be speculated that the high thermodynamic stability of alternative structures, may allow spontaneous folding *in vivo*, possibly mediated by subtle balance between cation concentrations (Darlow and Leach, 1998b), without the requirement for protein-mediated energy-dependent strand dissociation, or by simple DNA breathing.

It is not unreasonable to assume that being unable to extend the double-stranded trinucleotide sequence, which could imply breaking intrastrand base-pairs, and thereby eliminating any stable secondary structure, MMR proteins are left with two options: they can either remove the alternative structure, by DNA excision and re-ligation, or synthesise a complementary DNA sequence opposite to the loop. In the absence of cell division, there is no strand discrimination signal to direct MMR proteins onto a particular DNA strand. Under these circumstances, it may seem logical that the MMR system may opt for the incorporation of an extra DNA sequence opposite to the alternative structure, since excision might result in the loss of an important piece of genetic information (Figure 9.1). Successive rounds of error-prone MMR, would lead to a time-dependent accumulation of repeat size changes in post-mitotic cells, such as the skeletal muscle in DM1, and the striatum in HD, in the absence of cell division and DNA replication. Moreover, on

the basis of the proposed model, the possible saturation of the MMR system by the requirements for rapid cell division remains as an attractive hypothesis to explain the relative repeat stabilisation in rapidly proliferating tissues (*e.g.* blood) and cell lines (*e.g.* *Dmt-D* lungs cell cultures). Although highly speculative and lack of direct experimental data to support the crucial expansion bias at the MMR level that is the basis of this model, this hypothetical mechanism of trinucleotide instability could not only explain repeat length changes in post-mitotic tissues, but it is also consistent with most of the reported data based on mammalian systems, particularly transgenic mouse models.

### **9.1.3. Dynamic equilibrium between MMR heteroduplexes and trinucleotide repeat mutation**

It has recently been reported that competition between Msh3 and Msh6 proteins for binding to Msh2 mediates the outcome of distinct MMR deficiencies in the levels of trinucleotide repeat somatic mosaicism (van Den Broek *et al.*, 2002). It is conceivable that different MMR heterodimers exhibit different roles in recognition and/or processing of non-orthodox DNA structures formed within trinucleotide repetitive sequences, dictating the effect of MMR mutation on repeat dynamics. The data reported for a knock-in DM1 mouse model, suggests that only MMR complexes containing Msh2 and Msh3, involved in the repair of insertion/deletion loops within microsatellite sequences (Figure 1.3), are specifically responsible for the triplet repeat expansion mutations. In contrast, MMR complexes consisting of Msh2 and Msh6, being specifically targeted for single-base mismatches (Figure 1.3), are not involved in the mutation mechanism. Indeed, *Msh6* deletion results in a significant increase of somatic mosaicism (van Den Broek *et al.*, 2002). Since MutS $\alpha$  heterodimers (Msh2-Msh6) are in large excess to MutS $\beta$  heterodimers (Msh2-Msh3) or free Msh2 in wild-type cells (de Wind *et al.*, 1999), it appears logical to assume that in the absence of Msh6, free Msh2 is recruited by Msh3 and targeted to loop/hairpin repair, which through DNA mis-incorporation opposite to the secondary structure, may result in repeat gain. In a similar way, the recruitment of Msh2-Msh6 heteroduplexes by UV- and ROS-induced DNA lesions (Mu *et al.*, 1997; Ni *et al.*, 1999; Wang *et al.*, 1999a), may explain the lower levels of repeat expansion detected in culture following exposure to ROS-generating compounds (Chapter 5) or ultraviolet (UV) light (Chapter 6). Therefore, to be more precise, the MutS homologues represented in Figure 9.1 should only include MSH2 and MSH3.

Given the excess of MSH2-MSH6 complexes, relative to MSH2-MSH3, and the opposing effects of the two protein heteroduplexes on the dynamics of trinucleotide repeats, it appears logical to speculate that a competition between MSH3 and MSH6 to bind MSH2, may be a key step in the mutation process. The ratio between the two MutS heteroduplexes is likely to be an important determinant of the alternative pathways operating on mutation intermediates, thereby controlling the expansion rates of trinucleotide repeat sequences. This hypothesis could be tested by treating *Dmt-D* mouse cells with methotrexate. Methotrexate induces a genomic amplification event of the *dihydrofolate reductase* gene, co-amplifying the *Msh3* gene as well (Drummond *et al.*, 1997). Due to overproduction of Msh3 and heterodimer formation of this protein with virtually

all the nuclear Msh2, higher levels of expansion-biased repeat instability would therefore be expected.

It is noteworthy that the reduction of somatic trinucleotide repeat instability in the absence of a functional *Pms2* gene revealed by SP-PCR studies (Chapter 8), although highly significant, is not as dramatic as the consequences of *Msh2* and *Msh3* deletion in other mouse models (Manley *et al.*, 1999b; van Den Broek *et al.*, 2002), revealed by high DNA template levels PCR amplification, using fluorescent primers, automated fragment analysis and GeneScan software. These apparently discrepant results may be, at least partially, explained by the different methods used to assess trinucleotide repeat variability. However, the different extent in which MMR mutations affect trinucleotide repeat stability, most likely result from the overlapping functions of *Pms2* and *Mlh1* in the repair of insertion/deletion loops (Lipkin *et al.*, 2000). In the total absence of *Pms2* protein, *Mlh1* can still form a functional MutL $\beta$  heterodimer with *Mlh3*, and contribute to the expansion of trinucleotide repeats.

Overexpression of MMR genes, the use of dominant and/or anti-sense technology, could not only provide further insight into the actual MMR genes directly involved in the expansion mechanism of trinucleotide repeats, but also help to establish the importance of the dynamic equilibrium between MutS and MutL homologues in the expansion-bias mutation of triplet repeat sequences.

#### **9.1.4. Involvement of NER, BER and recombination events in triplet repeat mutation**

In addition to the direct involvement of the MMR pathway in the mechanism of trinucleotide repeat mutation, other repair mechanisms may also be involved. The exposure of cultured cells to genotoxic agents that mainly activate nucleotide or base excision repair (NER and BER), such as ROS-generating compounds (Chapter 5) and UV light (Chapter 6), was shown to alter the dynamics of expanded trinucleotide repeats in culture. Whether the outcome of such treatments reflects a direct effect on the mutation process, or a selection of cells with enhanced resistance to the specific genotoxic agents, remains unclear. However, the possibility exists that activation of NER and BER may lead to the alternative processing of mutation intermediates, resulting in an altered mutation profile.

In addition to MMR, NER, BER and more specialised systems to deal with specific DNA lesions, there is also a need for enzymatic complexes to maintain the structural integrity of DNA, which might be broken and give rise to double strand breaks. Double strand breaks can arise in many ways, such as failure of topoisomerase II to complete its cleavage and rejoining cycle, mechanical rupture of tangled sister chromatids during mitosis, or by ionising radiation. However, the major source of double strand breaks comes from the process of DNA replication itself. These breaks probably arise from events that occur at stalled replication forks. There are two major types of double strand break repair: non-homologous end-joining, in which ends are re-joined, often with the loss of a few base pairs of sequence; homologous recombination, which can lead to accurate

repair of the double strand break by using a homologous chromosome, especially a sister chromatid, as the template to "fill-in" the gap (Jackson, 2002). Regardless of whether it is theoretically possible, there is as yet no direct evidence on the extent to which, if at all, recombination processes are responsible for trinucleotide repeat expansions in mammalian cells. If recombination between homologous chromosomes were the major source of repeat instability in somatic cells, mice carrying two expanded CAG•CTG repeat sequences would exhibit greater levels of somatic mosaicism. This appears not to be the case in a knock-in mouse model of HD (L. Kennedy and P.F. Shelbourne, personal communication). Furthermore, *Dmt-D* hemizygous mice exhibit dramatic somatic mosaicism, although they only carry a single *Dmt* transgene copy in their genome (Fortune *et al.*, 2000). Alternatively, if sister chromatid exchange were the driving force of repeat mutation, the process would be dependent on cell division. However, mounting evidence suggests this is not the case either (Chapter 7). In addition, double strand break-mediated recombination repair cannot account for the full mutation spectrum observed in human disease, because double strand break at CAG•CTG repeats in yeast are observed only above 130 repeat units (Freudenreich *et al.*, 1998). Nevertheless, some of the repeat length mutations observed in *Dmt-D* cultured cells might have been mediated by recombination-based events, given that several treatments might have induced double strand DNA breaks, which in turn might have triggered a recombination-based repair mechanism.

In summary, despite the major role of the MMR pathway in the expansion of trinucleotide repeats, polymerase slippage occurring during both DNA repair or replication synthesis, NER and BER pathways are also likely to be involved in the mutation mechanism. Since MMR proteins are also involved in NER and BER (Buermeyer *et al.*, 1999), there is the possibility that the central players in all these co-operating mechanisms might be the same: MMR enzymes.

### **9.1.5. Central involvement of DNA topology in the dynamics of expanded trinucleotide repeats**

In addition to MMR proteins, the results presented in Chapter 4 brought new players into the picture. Enzymes involved in DNA topology, such as topoisomerases, may also be involved in triplet repeat metabolism. The double helical nature of DNA creates topological problems that must be solved in order to allow the proper transmission of the genetic information. Two steps are required for this process: disruption of hydrogen bonds between the two DNA strands, performed by specialised DNA helicases; elimination of all the topological links between the two strands, performed by DNA topoisomerases (Duguet, 1997). The cooperation between helicases and topoisomerases is likely to be extended to many aspects of DNA metabolism, including progression of the DNA replication fork, segregation of newly replicated chromosomes, disruption of nucleosomal structure (especially during transcription), DNA super coiling and finally DNA recombination, repair and genome stability. The role of helicases seems to be so important, that as much as 1% of eukaryotic genes may indeed encode DNA and/or RNA helicase enzymes. In addition, helicase deficiencies have been associated with human genetic disorders featuring

inherent genomic instability, such as elevated frequency of homologous and illegitimate recombination, chromosomal deletions and aberrations (Mohaghegh and Hickson, 2001). The role of DNA topological enzymes in maintaining genome integrity has thus been well documented. Because the duplex helix structure associated with triplet repeat DNA sequence is so anomalous compared with the mixed-sequence DNA found in most of the genome, the interaction of helicases and topoisomerases with trinucleotide repetitive tracts may be quite distinctive. Indeed, some helicases have been reported to interact and unwind tetraplex structures formed by single-stranded CGG repeat sequences (Fry and Loeb, 1999). Nevertheless, their involvement in the molecular mechanisms of trinucleotide repeat mutation has never been described.

Further evidence supporting the involvement of topoisomerase in the dynamics of expanded trinucleotide repeats could be achieved by testing other topoisomerase inhibitors, or by transforming *Dmt-D* cell lines with dominant negatives of topoisomerase or helicase enzymes, and monitoring the repeat size by sensitive SP-PCR techniques. Additionally, cellular proteins capable of binding CAG•CTG repeats could be purified by binding affinity procedures. Subsequent two-dimensional gel electrophoresis could lead to the identification of proteins that may interact with expanded CAG•CTG DNA repeats, and possibly mediate trinucleotide repeat metabolism *in vivo*.

## 9.2. On the molecular bases of tissue- and cell-specific somatic mosaicism

The patterns of tissue-specific somatic mosaicism exhibited by transgenic mice carrying expanded CAG•CTG repeats share striking similarities. Most notably, the transgenic repeats appear particularly unstable in kidney, liver and striatum, while relatively low levels of instability are detected in cerebellum and spleen in all different mouse models (Fortune *et al.*, 2000; Lia *et al.*, 1998; Mangiarini *et al.*, 1997; Seznec *et al.*, 2000; van Den Broek *et al.*, 2002). The fact that the patterns of tissue-specific somatic mosaicism are broadly similar between independent transgenic lines suggests a major role for tissue-specific *trans*-acting modifiers of somatic trinucleotide repeat instability. Locus-dependent *cis* factors may, however, contribute to the dissimilarities between mutation profiles in different mouse models, as proposed for trinucleotide instability in the human germline (Brock *et al.*, 1999).

According to the model presented (Section 9.1.2, Figure 9.1), tissue-specific repeat instability might be partially explained by different expression levels of the components of MMR in different tissues. Either because MMR naturally shows tissue-specific expression profiles (or even efficiency profiles), or because some tissues are more prone to DNA damage, which will subsequently activate different DNA repair systems.

Tissue-specific functions are predicted for MMR genes, according to the distinct tumour spectra observed in MMR-deficient mice (Section 1.6.3.1 and Table 1.4). It is also conceivable that the same differences may account for the tissue-specific patterns of trinucleotide repeat instability in somatic cells. In addition, tissue-specific expression levels have already been reported for some

MMR genes (Lipkin *et al.*, 2000; Prolla *et al.*, 1998), together with cell type-specific alternative splicing forms for the human *MSH2* gene (Clarke *et al.*, 2000). These findings may well account for the characteristic tissue-specific patterns of trinucleotide repeat somatic mosaicism described in mice and humans.

Quantitative *in vivo* assays of DNA MMR activity has allowed the identification of MMR efficiency polymorphisms, caused by amino acid replacements that do not inactivate MMR, but result in lower efficiency of DNA repair (Ellison *et al.*, 2001). Similarly, missense mutations in yeast *msh2*, *msh3*, *mlh1*, and *pms1* MMR genes have been shown to cause phenotypes that are different from those of null mutations (Sia *et al.*, 2001). These observations raise the possibility that differences in the efficiency of DNA MMR, which have no detectable phenotypic outcome, exist between individuals in the population due to common polymorphisms. Together with putative haplotype-associated *cis*-acting factors, such differences may explain why different individuals who have inherited similar size repeat sequences, exhibit different levels of instability, and possibly distinct clinical pictures in respect to disease severity.

Spontaneous mutations have been analysed at a variety of loci *in vitro*. The spontaneous mutation spectrum depends heavily, not only on the genetic marker and its chromosomal location, but also on the cell type (Glickman *et al.*, 1994; Gossen *et al.*, 1993). In addition, levels of oxidative DNA damage appears to be broadly related to metabolic rate in a number of mammalian species (Ames *et al.*, 1993). Oxidised DNA bases are indeed present at high levels in a wide range of human and rodent tissues and the variation in adduct levels correlates roughly with metabolic rates (Marnett, 2000). Exposure of *Dmt-D* cells to acute levels of oxidative stress resulted in decreased rates of trinucleotide repeat expansion (Chapter 5), probably mediated by cell selection in culture, alternative processing of mutation intermediates or inhibition of MMR proteins, via protein oxidation (Chang *et al.*, 2002). Although it is unlikely that the severity of the treatments described recreates what happens *in vivo*, where a slow build up of oxidants over time is more likely to occur, the results presented in Chapter 4 strongly suggest that ROS may have an impact on repeat metabolism. It is conceivable that different tissues and cell types produce different endogenous mutagens, thereby activating distinct DNA repair pathways (MMR, NER and/or BER) to different extents, accounting for the tissue-specific triplet repeat mutation profiles detected, not only in humans, but also in mice.

Subtle differences in the cellular environment between different tissues and cell types, such as ion levels, could influence the proclivity of trinucleotide repeat tracts to adopt alternative structures, and therefore account for the differences in the expansion rates between tissues and cell types (Darlow and Leach, 1998b).

In summary, an intricate interplay of a variety of factors, involving DNA repair pathways, endogenous and exogenous genotoxic agents, and different DNA structural dynamics, may establish the central molecular bases of the tissue- and/or cell-specific nature of somatic mosaicism. In order to pinpoint the actual cell types exhibiting the highest expansion rates, sensitive SP-PCR analysis could be coupled to laser microdissection techniques or fluorescent activated cell sorting (FACS) methods. These assays should be particularly useful for tissues showing multimodal

mutation profiles, such as liver and kidney. In addition, to further test the involvement of oxidative stress as modifier of repeat dynamics, measurement of reactive oxygen levels within cells, either in culture or *in vivo* (following enzymatic dissociation), by 2,4-dichlorofluorescein (H<sub>2</sub>DCF) oxidation-based methods (Jakubowski and Bartosz, 2000), might prove extremely informative. The association of SP-PCR techniques with H<sub>2</sub>DCF fluorescence methods, could possibly reveal a conclusive relationship between levels of oxidative stress and tissue- or cell-specific somatic mosaicism.

### 9.3. Future perspectives

*Dmt-D* mouse cell cultures appear to be a suitable and powerful model to investigate the dynamics of expanded trinucleotide repeat dynamics. However spontaneous immortalised mouse cell lines may carry many accumulated mutations and abnormalities that may interfere with DNA metabolism, and directly affect the stability of triplet repeats. Therefore, the use of mouse embryonic fibroblasts, or primary cultures, might be more appropriate. However, these cells may not live long enough, or they may undergo spontaneous immortalisation too soon, thereby reducing the time window to perform experiments similar to those reported in previous chapters. Nevertheless, mouse embryonic fibroblasts should be of great interest to assess the effects of proliferation inhibition by cell-to-cell contact, which cannot be achieved with immortalised mouse cell lines without exposure to drugs. It is also noteworthy that the selective sweeps described in Chapter 3, which correspond to unpredicted periods of reduced repeat length variability, may still influence the outcome of drug treatments on the dynamics of expanded CAG•CTG repeats, even when multiple replicates are established. The use of naturally arrested cells by contact inhibition, could help to overcome this limitation of the *Dmt-D* cell culture model, and create new avenues to assess the effect of potential therapeutical chemicals on post-mitotic cells, which are usually the most affected cells in these disorders.

Nevertheless, selective sweeps may also prove informative. Microarray expression analysis, may lead to the identification of gene mutations or altered gene expression profiles associated with lower degrees of repeat length heterogeneity, thereby providing further insight into the genetic bases of triplet repeat mutation.

#### 9.3.1. Novel therapeutical routes

Whilst it is unlikely that most of the specific compounds tested here, with a few exceptions, would have therapeutic utility in humans, the results presented may be considered as a proof of principle and a basis for more extended screens. Some of these effects reported might be mediated by cell selection in rapidly dividing cultures. Nonetheless, drugs that resulted in decreased repeat lengths through cell selection could still have therapeutic benefits. As a step forward, towards the development of novel therapeutical routes, and towards a better understanding

of trinucleotide repeat dynamics in mammalian systems, promising agents identified *in vitro*, should have their utility assessed *in vivo*, in *Dmt-D* transgenic mice, in which the unstable DNA phenotype is replicated. The assessment of the effects of novobiocin, and possibly other topoisomerase inhibitors, as well as aspirin on the dynamics of the transgene in these animals should be of great interest. If confirmed, these findings would render “chemogenotherapy” as a possible route to treat or prevent disease progression, as a means of modifying the nature of endogenous DNA sequences by the application of exogenous chemicals. The possibility is very attractive since it might be achievable with small chemical drugs, for which delivery methods would be less problematic.

The possibility of a presymptomatic diagnosis and the dramatic nature of the disease have prompted major efforts to develop early treatments to block the effect of the mutation before the symptoms show up. If a preventive treatment at the DNA levels is possible, it should therefore be administered to all individuals with the mutation as early as possible, without waiting for the appearance of symptoms. This would obviously imply a mutation screening for trinucleotide repeat expansions in the general asymptomatic population, which would raise ethical questions, which must be acknowledged and debated.

# References

- Abel, A., Walcott, J., Woods, J., Duda, J., and Merry, D. E. (2001). Expression of expanded repeat androgen receptor produces neurologic disease in transgenic mice. *Hum Mol Genet* **10**: 107-116.
- Abraham, R. T. (2001). Cell cycle checkpoint signaling through the ATM and ATR kinases. *Genes Dev* **15**: 2177-2196.
- Aburatani, H., Hippo, Y., Ishida, T., Takashima, R., Matsuba, C., Kodama, T., Takao, M., Yasui, A., Yamamoto, K., and Asano, M. (1997). Cloning and characterization of mammalian 8-hydroxyguanine-specific DNA glycosylase/apurinic, apyrimidinic lyase, a functional mutM homologue. *Cancer Res* **57**: 2151-2156.
- Adachi, H., Kume, A., Li, M., Nakagomi, Y., Niwa, H., Do, J., Sang, C., Kobayashi, Y., Doyu, M., and Sobue, G. (2001). Transgenic mice with an expanded CAG repeat controlled by the human AR promoter show polyglutamine nuclear inclusions and neuronal dysfunction without neuronal cell death. *Hum Mol Genet* **10**: 1039-1048.
- Allen, R. G., and Tresini, M. (2000). Oxidative stress and gene regulation. *Free Radic Biol Med* **28**: 463-499.
- Allingham-Hawkins, D. J., Babul-Hirji, R., Chitayat, D., Holden, J. J., Yang, K. T., Lee, C., Hudson, R., Gorwill, H., Nolin, S. L., Glicksman, A., *et al.* (1999). Fragile X premutation is a significant risk factor for premature ovarian failure: the International Collaborative POF in Fragile X study- -preliminary data. *Am J Med Genet* **83**: 322-325.
- Alwazzan, M., Newman, E., Hamshere, M. G., and Brook, J. D. (1999). Myotonic dystrophy is associated with a reduced level of RNA from the DMWD allele adjacent to the expanded repeat. *Hum Mol Genet* **8**: 1491-1497.
- Amacher, D. E., and Turner, G. N. (1987). The mutagenicity of 5-azacytidine and other inhibitors of replicative DNA synthesis in the L5178Y mouse lymphoma cell. *Mutat Res* **176**: 123-131.
- Amack, J. D., Paguio, A. P., and Mahadevan, M. S. (1999). Cis and trans effects of the myotonic dystrophy (DM) mutation on a cell culture. *Human Molecular Genetics* **8**: 1975-1984.
- Ames, B. N., Shigenaga, M. K., and Hagen, T. M. (1993). Oxidants, antioxidants, and the degenerative diseases of aging. *Proc Natl Acad Sci U S A* **90**: 7915-7922.
- Andoh, T., and Ishida, R. (1998). Catalytic inhibitors of DNA topoisomerase II. *Biochim Biophys Acta* **1400**: 155-171.
- Andrew, S. E., Xu, X. S., Baross-Francis, A., Narayanan, L., Milhausen, K., Liskay, R. M., Jirik, F. R., and Glazer, P. M. (2000). Mutagenesis in PMS2- and MSH2-deficient mice indicates differential protection from transversions and frameshifts. *Carcinogenesis* **21**: 1291-1295.
- Anvret, M., Ahlberg, G., Grandell, U., Hedberg, B., Johnson, K., and Edstrom, L. (1993). Larger expansions of the CTG repeat in muscle compared to lymphocytes from patients with myotonic dystrophy. *Hum Mol Genet* **2**: 1397-1400.
- Aquilina, G., and Bignami, M. (2001). Mismatch repair in correction of replication errors and processing of DNA damage. *J Cell Physiol* **187**: 145-154.
- Arai, K., Morishita, K., Shinmura, K., Kohno, T., Kim, S. R., Nohmi, T., Taniwaki, M., Ohwada, S., and Yokota, J. (1997). Cloning of a human homolog of the yeast OGG1 gene that is involved in the repair of oxidative DNA damage. *Oncogene* **14**: 2857-2861.
- Arango, H. A., Icely, S., Roberts, W. S., Cavanagh, D., and Becker, J. L. (2001). Aspirin effects on endometrial cancer cell growth. *Obstet Gynecol* **97**: 423-427.
- Artero, R., Prokop, A., Paricio, N., Begemann, G., Pueyo, I., Mlodzik, M., Perez-Alonso, M., and Baylies, M. K. (1998). The muscleblind gene participates in the organization of Z-bands and epidermal attachments of *Drosophila* muscles and is regulated by Dmef2. *Dev Biol* **195**: 131-143.
- Ashizawa, T., Dubel, J. R., and Harati, Y. (1993). Somatic instability of CTG repeat in myotonic dystrophy. *Neurology* **43**: 2674-2678.
- Ashizawa, T., Monckton, D. G., Vaishnav, S., Patel, B. J., Voskova, A., and Caskey, C. T. (1996). Instability of the expanded (CTG)<sub>n</sub> repeats in the myotonin protein kinase gene in cultured lymphoblastoid cell lines from patients with myotonic dystrophy. *Genomics* **36**: 47-53.
- Aslanidis, C., Jansen, G., Amemiya, C., Shutler, G., Mahadevan, M., Tsilfidis, C., Chen, C., Alleman, J., Wormskamp, N. G., Vooijs, M., and *et al.* (1992). Cloning of the essential myotonic dystrophy region and mapping of the putative defect. *Nature* **355**: 548-551.

- Avadhani, N. G., and Rutman, R. J. (1975). Differential effects of ethidium bromide on cytoplasmic and mitochondrial protein synthesis. *FEBS Lett* **50**: 303-305.
- Bacolla, A., Bowater, R. P., and Wells, R. D. (1998). Systems for the study of genetic instabilities: *Escherichia coli*. In *Genetic Instabilities and Hereditary Neurological Diseases*, R. D. Wells, and S. T. Warren, eds. (San Diego, Academic Press).
- Bacolla, A., Gellibolian, R., Shimizu, M., Amirhaeri, S., Kang, S., Ohshima, K., Larson, J. E., Harvey, S. C., Stollar, B. D., and Wells, R. D. (1997). Flexible DNA: genetically unstable CTG.CAG and CGG.CCG from human hereditary neuromuscular disease genes. *J Biol Chem* **272**: 16783-16792.
- Bader, S., Walker, M., Hendrich, B., Bird, A., Bird, C., Hooper, M., and Wyllie, A. (1999). Somatic frameshift mutations in the MBD4 gene of sporadic colon cancers with mismatch repair deficiency. *Oncogene* **18**: 8044-8047.
- Baker, S. M., Bronner, C. E., Zhang, L., Plug, A. W., Robatzek, M., Warren, G., Elliott, E. A., Yu, J., Ashley, T., Arnheim, N., and et al. (1995). Male mice defective in the DNA mismatch repair gene PMS2 exhibit abnormal chromosome synapsis in meiosis. *Cell* **82**: 309-319.
- Baldi, P., Brunak, S., Chauvin, Y., and Pedersen, A. G. (1999). Structural basis for triplet repeat disorders: a computational analysis. *Bioinformatics* **15**: 918-929.
- Bar-Ziv, R., and Libchaber, A. (2001). Effects of DNA sequence and structure on binding of RecA to single-stranded DNA. *Proc Natl Acad Sci U S A* **98**: 9068-9073.
- Barnetson, R., Jass, J., Tse, R., Eckstein, R., Robinson, B., and Schnitzler, M. (2000). Mutations associated with microsatellite unstable colorectal carcinomas exhibit widespread intratumoral heterogeneity. *Genes Chromosomes Cancer* **29**: 130-136.
- Bartek, J., and Lukas, J. (2001). Mammalian G1- and S-phase checkpoints in response to DNA damage. *Curr Opin Cell Biol* **13**: 738-747.
- Bates, G. P., and Davies, S. W. (1997). Transgenic mouse models of neurodegenerative disease caused by CAG/polyglutamine expansions. *Mol Med Today* **3**: 508-515.
- Bearzatto, A., Szadkowski, M., Macpherson, P., Jiricny, J., and Karran, P. (2000). Epigenetic regulation of the MGMT and hMSH6 DNA repair genes in cells resistant to methylating agents. *Cancer Res* **60**: 3262-3270.
- Becher, M. W., Kotzuc, J. A., Sharp, A. H., Davies, S. W., Bates, G. P., Price, D. L., and Ross, C. A. (1998). Intranuclear neuronal inclusions in Huntington's disease and dentatorubral and pallidolusian atrophy: correlation between the density of inclusions and IT15 CAG triplet repeat length. *Neurobiol Dis* **4**: 387-397.
- Beckman, K. B., and Ames, B. N. (1997). Oxidative decay of DNA. *J Biol Chem* **272**: 19633-19636.
- Begemann, G., Paricio, N., Artero, R., Kiss, I., Perez-Alonso, M., and Mlodzik, M. (1997). muscleblind, a gene required for photoreceptor differentiation in *Drosophila*, encodes novel nuclear Cys3His-type zinc-finger-containing proteins. *Development* **124**: 4321-4331.
- Bellacosa, A. (2001). Role of MED1 (MBD4) Gene in DNA repair and human cancer. *J Cell Physiol* **187**: 137-144.
- Bellacosa, A., Cicchillitti, L., Schepis, F., Riccio, A., Yeung, A. T., Matsumoto, Y., Golemis, E. A., Genuardi, M., and Neri, G. (1999). MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. *Proc Natl Acad Sci U S A* **96**: 3969-3974.
- Belyaev, I. Y., Eriksson, S., Nygren, J., Torudd, J., and Harms-Ringdahl, M. (1999). Effects of ethidium bromide on DNA loop organisation in human lymphocytes measured by anomalous viscosity time dependence and single cell gel electrophoresis. *Biochim Biophys Acta* **1428**: 348-356.
- Bender, C. M., Pao, M. M., and Jones, P. A. (1998). Inhibition of DNA methylation by 5-aza-2'-deoxycytidine suppresses the growth of human tumor cell lines. *Cancer Res* **58**: 95-101.
- Benders, A. A., Groenen, P. J., Oerlemans, F. T., Veerkamp, J. H., and Wieringa, B. (1997). Myotonic dystrophy protein kinase is involved in the modulation of the Ca<sup>2+</sup> homeostasis in skeletal muscle cells. *J Clin Invest* **100**: 1440-1447.
- Bennett, R. L., and Holloman, W. K. (2001). A RecA homologue in *Ustilago maydis* that is distinct and evolutionarily distant from Rad51 actively promotes DNA pairing reactions in the absence of auxiliary factors. *Biochemistry* **40**: 2942-2953.
- Berger, J. M. (1998). Structure of DNA topoisomerases. *Biochim Biophys Acta* **1400**: 3-18.
- Bertrand, P., Tishkoff, D. X., Filosi, N., Dasgupta, R., and Kolodner, R. D. (1998). Physical interaction between components of DNA mismatch repair and nucleotide excision repair. *Proc Natl Acad Sci U S A* **95**: 14278-14283.

- Bhagwati, S., Ghatpande, A., and Leung, B. (1996). Normal levels of DM RNA and myotonin protein kinase in skeletal muscle from adult myotonic dystrophy (DM) patients. *Biochim Biophys Acta* **1317**: 155-157.
- Bhagwati, S., Shafiq, S. A., and Xu, W. (1999). (CTG)<sub>n</sub> repeats markedly inhibit differentiation of the C1C12 myoblast cell line: implications for congenital myotonic dystrophy. *Biochimica et Biophysica Acta* **1453**: 221-229.
- Bhui-Kaur, A., Goodman, M. F., and Tower, J. (1998). DNA mismatch repair catalyzed by extracts of mitotic, postmitotic, and senescent *Drosophila* tissues and involvement of mei-9 gene function for full activity. *Mol Cell Biol* **18**: 1436-1443.
- Bidichandani, S. I., Ashizawa, T., and Patel, P. I. (1998). The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. *Am J Hum Genet* **62**: 111-121.
- Bidichandani, S. I., Purandare, S. M., Taylor, E. E., Gumin, G., Machkhas, H., Harati, Y., Gibbs, R. A., Ashizawa, T., and Patel, P. I. (1999). Somatic sequence variation at the Friedreich ataxia locus includes complete contraction of the expanded GAA triplet repeat, significant length variation in serially passaged lymphoblasts and enhanced mutagenesis in the flanking sequence. *Hum Mol Genet* **8**: 2425-2436.
- Bigioni, M., Zunino, F., Tinelli, S., Austin, C. A., Willmore, E., and Capranico, G. (1996). Position-specific effects of base mismatch on mammalian topoisomerase II DNA cleaving activity. *Biochemistry* **35**: 153-159.
- Bird, A. P., and Wolffe, A. P. (1999). Methylation-induced repression - belts, braces, and chromatin. *Cell* **99**: 451-454.
- Bjoras, M., Luna, L., Johnsen, B., Hoff, E., Haug, T., Rognes, T., and Seeberg, E. (1997). Opposite base-dependent reactions of a human base excision repair enzyme on DNA containing 7,8-dihydro-8-oxoguanine and abasic sites. *Embo J* **16**: 6314-6322.
- Blasina, A., Price, B. D., Turenne, G. A., and McGowan, C. H. (1999). Caffeine inhibits the checkpoint kinase ATM. *Curr Biol* **9**: 1135-1138.
- Bogdanov, M. B., Andreassen, O. A., Dedeoglu, A., Ferrante, R. J., and Beal, M. F. (2001). Increased oxidative damage to DNA in a transgenic mouse model of Huntington's disease. *J Neurochem* **79**: 1246-1249.
- Bontekoe, C. J., Bakker, C. E., Nieuwenhuizen, I. M., van der Linde, H., Lans, H., de Lange, D., Hirst, M. C., and Oostra, B. A. (2001). Instability of a (CGG)<sub>98</sub> repeat in the *Fmr1* promoter. *Hum Mol Genet* **10**: 1693-1699.
- Boucher, C. A., King, S. K., Carey, N., Krahe, R., Winchester, C. L., Rahman, S., Creavin, T., Meghji, P., Bailey, M. E., Chartier, F. L., and et al. (1995). A novel homeodomain-encoding gene is associated with a large CpG island interrupted by the myotonic dystrophy unstable (CTG)<sub>n</sub> repeat. *Hum Mol Genet* **4**: 1919-1925.
- Bowater, R. P., Jaworski, A., Larson, J. E., Parniewski, P., and Wells, R. D. (1997). Transcription increases the deletion frequency of long CTG.CAG triplet repeats from plasmids in *Escherichia coli*. *Nucleic Acids Res* **25**: 2861-2868.
- Bowater, R. P., Rosche, W. A., Jaworski, A., Sinden, R. R., and Wells, R. D. (1996). Relationship between *Escherichia coli* growth and deletions of CTG.CAG triplet repeats in plasmids. *J Mol Biol* **264**: 82-96.
- Brevik, J., and Gaudernack, G. (1999). Genomic instability, DNA methylation, and natural selection in colorectal carcinogenesis. *Semin Cancer Biol* **9**: 245-254.
- Breschel, T. S., McInnis, M. G., Margolis, R. L., Sirugo, G., Corneliussen, B., Simpson, S. G., McMahon, F. J., MacKinnon, D. F., Xu, J. F., Pleasant, N., et al. (1997). A novel, heritable, expanding CTG repeat in an intron of the *SEF2-1* gene on chromosome 18q21.1. *Hum Mol Genet* **6**: 1855-1863.
- Brezden, C. B., and Rauth, A. M. (1996). Differential cell death in immortalized and non-immortalized cells at confluency. *Oncogene* **12**: 201-206.
- Brock, G. J., Anderson, N. H., and Monckton, D. G. (1999). Cis-acting modifiers of expanded CAG/CTG triplet repeat expandability: associations with flanking GC content and proximity to CpG islands. *Hum Mol Genet* **8**: 1061-1067.
- Brook, J. D., McCurrach, M. E., Harley, H. G., Buckler, A. J., Church, D., Aburatani, H., Hunter, K., Stanton, V. P., Thirion, J. P., Hudson, T., and et al. (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* **69**: 385.
- Brown, W. T., Houck, G. E., Jr., Ding, X., Zhong, N., Nolin, S., Glicksman, A., Dobkin, C., and Jenkins, E. C. (1996). Reverse mutations in the fragile X syndrome. *Am J Med Genet* **64**: 287-292.

- Brukner, I., Belmaaza, A., and Chartrand, P. (1997). Differential behavior of curved DNA upon untwisting. *Proc Natl Acad Sci U S A* **94**: 403-406.
- 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, B. A., Wieringa, B., and et al. (1993). Influence of sex 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.
- Buermeyer, A. B., Deschenes, S. M., Baker, S. M., and Liskay, R. M. (1999). Mammalian DNA mismatch repair. *Annu Rev Genet* **33**: 533-564.
- Burcham, P. C. (1999). Internal hazards: baseline DNA damage by endogenous products of normal metabolism. *Mutat Res* **443**: 11-36.
- Burden, D. A., and Osheroff, N. (1998). Mechanism of action of eukaryotic topoisomerase II and drugs targeted to the enzyme. *Biochim Biophys Acta* **1400**: 139-154.
- 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., and et al. (1992). Detection of an unstable fragment of DNA specific to individuals with myotonic dystrophy. *Nature* **355**: 547-548.
- Campuzano, V., Montermini, L., Lutz, Y., Cova, L., Hindelang, C., Jiralerspong, S., Trotter, Y., Kish, S. J., Faucheux, B., Trouillas, P., et al. (1997). Frataxin is reduced in Friedreich ataxia patients and is associated with mitochondrial membranes. *Hum Mol Genet* **6**: 1771-1780.
- Campuzano, V., Montermini, L., Molto, M. D., Pianese, L., Cossee, M., Cavalcanti, F., Monros, E., Rodius, F., Duclos, F., Monticelli, A., and et al. (1996). Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science* **271**: 1423-1427.
- Cancel, G., Abbas, N., Stevanin, G., Durr, A., Chneiweiss, H., Neri, C., Duyckaerts, C., Penet, C., Cann, H. M., Agid, Y., and et al. (1995). Marked phenotypic heterogeneity associated with expansion of a CAG repeat sequence at the spinocerebellar ataxia 3/Machado-Joseph disease locus. *Am J Hum Genet* **57**: 809-816.
- Canova, C., Chevalier, G., Remy, S., Brachet, P., and Wion, D. (1998). Epigenetic control of programmed cell death: inhibition by 5-azacytidine of 1,25-dihydroxyvitamin D<sub>3</sub>-induced programmed cell death in C6.9 glioma cells. *Mech Ageing Dev* **101**: 153-166.
- Capranico, G., and Binaschi, M. (1998). DNA sequence selectivity of topoisomerases and topoisomerase poisons. *Biochim Biophys Acta* **1400**: 185-194.
- 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.
- Carbone, G. M., Catapano, C. V., and Fernandes, D. J. (2001). Imbalanced DNA synthesis induced by cytosine arabinoside and fludarabine in human leukemia cells. *Biochem Pharmacol* **62**: 101-110.
- Cascalho, M., Wong, J., Steinberg, C., and Wabl, M. (1998). Mismatch repair co-opted by hypermutation. *Science* **279**: 1207-1210.
- Castaneda, F., and Kinne, R. K. (1999). Effects of doxorubicin, mitomycin C, and ethanol on Hep-G2 cells in vitro. *J Cancer Res Clin Oncol* **125**: 1-8.
- Castano, E., Dalmau, M., Marti, M., Berrocal, F., Bartrons, R., and Gil, J. (1997). Inhibition of DNA synthesis by aspirin in Swiss 3T3 fibroblasts. *J Pharmacol Exp Ther* **280**: 366-372.
- Catapano, C. V., Chandler, K. B., and Fernandes, D. J. (1991). Inhibition of primer RNA formation in CCRF-CEM leukemia cells by fludarabine triphosphate. *Cancer Res* **51**: 1829-1835.
- Cha, J. H., Kosinski, C. M., Kerner, J. A., Alsdorf, S. A., Mangiarini, L., Davies, S. W., Penney, J. B., Bates, G. P., and Young, A. B. (1998). Altered brain neurotransmitter receptors in transgenic mice expressing a portion of an abnormal human huntington disease gene. *Proc Natl Acad Sci U S A* **95**: 6480-6485.
- Champoux, J. J. (2001). DNA topoisomerases: structure, function, and mechanism. *Annu Rev Biochem* **70**: 369-413.
- Chan, T. A., Morin, P. J., Vogelstein, B., and Kinzler, K. W. (1998). Mechanisms underlying nonsteroidal antiinflammatory drug-mediated apoptosis. *Proc Natl Acad Sci U S A* **95**: 681-686.
- Chang, C. L., Marra, G., Chauhan, D. P., Ha, H. T., Chang, D. K., Ricciardiello, L., Randolph, A., Carethers, J. M., and Boland, C. R. (2002). Oxidative stress inactivates the human DNA mismatch repair system. *Am J Physiol Cell Physiol* **283**: C148-154.

- Chastain, P. D., 2nd, Eichler, E. E., Kang, S., Nelson, D. L., Levene, S. D., and Sinden, R. R. (1995). Anomalous rapid electrophoretic mobility of DNA containing triplet repeats associated with human disease genes. *Biochemistry* **34**: 16125-16131.
- Chastain, P. D., and Sinden, R. R. (1998). CTG repeats associated with human genetic disease are inherently flexible. *J Mol Biol* **275**: 405-411.
- Chen, D., Walsh, K., and Wang, J. (2000). Regulation of cdk2 activity in endothelial cells that are inhibited from growth by cell contact. *Arterioscler Thromb Vasc Biol* **20**: 629-635.
- Chen, K. H., Yakes, F. M., Srivastava, D. K., Singhal, R. K., Sobol, R. W., Horton, J. K., Van Houten, B., and Wilson, S. H. (1998a). Up-regulation of base excision repair correlates with enhanced protection against a DNA damaging agent in mouse cell lines. *Nucleic Acids Res* **26**: 2001-2007.
- Chen, R. Z., Pettersson, U., Beard, C., Jackson-Grusby, L., and Jaenisch, R. (1998b). DNA hypomethylation leads to elevated mutation rates. *Nature* **395**: 89-93.
- Chen, X., Mariappan, S. V., Catasti, P., Ratliff, R., Moyzis, R. K., Laayoun, A., Smith, S. S., Bradbury, E. M., and Gupta, G. (1995). Hairpins are formed by the single DNA strands of the fragile X triplet repeats: structure and biological implications. *Proc Natl Acad Sci U S A* **92**: 5199-5203.
- Chiurazzi, P., Pomponi, M. G., Pietrobono, R., Bakker, C. E., Neri, G., and Oostra, B. A. (1999). Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the FMR1 gene. *Hum Mol Genet* **8**: 2317-2323.
- Chong, S. S., Almqvist, E., Telenius, H., LaTray, L., Nichol, K., Bourdelat-Parks, B., Goldberg, Y. P., Haddad, B. R., Richards, F., Sillence, D., *et al.* (1997). Contribution of DNA sequence and CAG size to mutation frequencies of intermediate alleles for Huntington disease: evidence from single sperm analyses. *Hum Mol Genet* **6**: 301-309.
- Chong, S. S., McCall, A. E., Cota, J., Subramony, S. H., Orr, H. T., Hughes, M. R., and Zoghbi, H. Y. (1995). Gametic and somatic tissue-specific heterogeneity of the expanded SCA1 CAG repeat in spinocerebellar ataxia type 1. *Nat Genet* **10**: 344-350.
- Choudhry, S., Mukerji, M., Srivastava, A. K., Jain, S., and Brahmachari, S. K. (2001). CAG repeat instability at SCA2 locus: anchoring CAA interruptions and linked single nucleotide polymorphisms. *Hum Mol Genet* **10**: 2437-2446.
- Chung, M. Y., Ranum, L. P., 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.
- Clarke, A. A., Philpott, N. J., Gordon-Smith, E. C., and Rutherford, T. R. (1997). The sensitivity of Fanconi anaemia group C cells to apoptosis induced by mitomycin C is due to oxygen radical generation, not DNA crosslinking. *Br J Haematol* **96**: 240-247.
- Clarke, L. A., Jordan, P., and Boavida, M. G. (2000). Cell type specificity in alternative splicing of the human mismatch repair gene hMSH2. *Eur J Hum Genet* **8**: 347-352.
- Cleary, J. D., Nichol, K., Wang, Y. H., and Pearson, C. E. (2002). Evidence of cis-acting factors in replication-mediated trinucleotide repeat instability in primate cells. *Nat Genet* **31**: 37-46.
- Cleaver, J. E. (1994). It was a very good year for DNA repair. *Cell* **76**: 1-4.
- Cleaver, J. E., Karplus, K., Kashani-Sabet, M., and Limoli, C. L. (2001). Nucleotide excision repair "a legacy of creativity". *Mutat Res* **485**: 23-36.
- Clikeman, J. A., Wheeler, S. L., and Nickoloff, J. A. (2001). Efficient incorporation of large (>2 kb) heterologies into heteroduplex DNA: Pms1/Msh2-dependent and -independent large loop mismatch repair in *Saccharomyces cerevisiae*. *Genetics* **157**: 1481-1491.
- Cobo, A., Martinez, J. M., Martorell, L., Baiget, M., and Johnson, K. (1993). Molecular diagnosis of homozygous myotonic dystrophy in two asymptomatic sisters. *Hum Mol Genet* **2**: 711-715.
- Coffee, B., Zhang, F., Warren, S. T., and Reines, D. (1999). Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. *Nat Genet* **22**: 98-101.
- Collins, A. R., and Horvathova, E. (2001). Oxidative DNA damage, antioxidants and DNA repair: applications of the comet assay. *Biochem Soc Trans* **29**: 337-341.
- Coogan, P. F., Rao, S. R., Rosenberg, L., Palmer, J. R., Strom, B. L., Zauber, A. G., Stolley, P. D., and Shapiro, S. (1999). The relationship of nonsteroidal anti-inflammatory drug use to the risk of breast cancer. *Prev Med* **29**: 72-76.
- Cossee, M., Schmitt, M., Campuzano, V., Reutenauer, L., Moutou, C., Mandel, J. L., and Koenig, M. (1997). Evolution of the Friedreich's ataxia trinucleotide repeat expansion: founder effect and premutations. *Proc Natl Acad Sci U S A* **94**: 7452-7457.

- Cramer, D. W., Harlow, B. L., Titus-Ernstoff, L., Bohlke, K., Welch, W. R., and Greenberg, E. R. (1998). Over-the-counter analgesics and risk of ovarian cancer. *Lancet* **351**: 104-107.
- Crawford, D. C., Wilson, B., and Sherman, S. L. (2000). Factors involved in the initial mutation of the fragile X CCG repeat as determined by sperm small pool PCR. *Hum Mol Genet* **9**: 2909-2918.
- Crawford, D. R., Schools, G. P., Salmon, S. L., and Davies, K. J. (1996). Hydrogen peroxide induces the expression of adapt15, a novel RNA associated with polysomes in hamster HA-1 cells. *Arch Biochem Biophys* **325**: 256-264.
- Croteau, D. L., and Bohr, V. A. (1997). Repair of oxidative damage to nuclear and mitochondrial DNA in mammalian cells. *J Biol Chem* **272**: 25409-25412.
- Cummings, C. J., Reinstein, E., Sun, Y., Antalffy, B., Jiang, Y., Ciechanover, A., Orr, H. T., Beaudet, A. L., and Zoghbi, H. Y. (1999). Mutation of the E6-AP ubiquitin ligase reduces nuclear inclusion frequency while accelerating polyglutamine-induced pathology in SCA1 mice. *Neuron* **24**: 879-892.
- Cummings, C. J., and Zoghbi, H. Y. (2000a). Fourteen and counting: unraveling trinucleotide repeat diseases. *Hum Mol Genet* **9**: 909-916.
- Cummings, C. J., and Zoghbi, H. Y. (2000b). Trinucleotide repeats: mechanisms and pathophysiology. *Annu Rev Genomics Hum Genet* **1**: 281-328.
- Dalton, T. P., Shertzer, H. G., and Puga, A. (1999). Regulation of gene expression by reactive oxygen. *Annu Rev Pharmacol Toxicol* **39**: 67-101.
- Darlow, J. M., and Leach, D. R. (1995). The effects of trinucleotide repeats found in human inherited disorders on palindrome inviability in *Escherichia coli* suggest hairpin folding preferences in vivo. *Genetics* **141**: 825-832.
- Darlow, J. M., and Leach, D. R. (1998a). Evidence for two preferred hairpin folding patterns in d(CGG).d(CCG) repeat tracts in vivo. *J Mol Biol* **275**: 17-23.
- Darlow, J. M., and Leach, D. R. (1998b). Secondary structures in d(CGG) and d(CCG) repeat tracts. *J Mol Biol* **275**: 3-16.
- David, G., Abbas, N., Stevanin, G., Durr, A., Yvert, G., Cancel, G., Weber, C., Imbert, G., Saudou, F., Antoniou, E., *et al.* (1997). Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nat Genet* **17**: 65-70.
- Davies, K. J. (1999). The broad spectrum of responses to oxidants in proliferating cells: a new paradigm for oxidative stress. *IUBMB Life* **48**: 41-47.
- Davies, S. W., Turmaine, M., Cozens, B. A., DiFiglia, M., Sharp, A. H., Ross, C. A., Scherzinger, E., Wanker, E. E., Mangiarini, L., and Bates, G. P. (1997). Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* **90**: 537-548.
- Davies, S. W., Turmaine, M., Cozens, B. A., Raza, A. S., Mahal, A., Mangiarini, L., and Bates, G. P. (1999). From neuronal inclusions to neurodegeneration: neuropathological investigation of a transgenic mouse model of Huntington's disease. *Philos Trans R Soc Lond B Biol Sci* **354**: 971-979.
- Davis, B. M., McCurrach, M. E., Taneja, K. L., Singer, R. H., and Housman, D. E. (1997). Expansion of a CUG trinucleotide repeat in the 3' untranslated region of myotonic dystrophy protein kinase transcripts results in nuclear retention of transcripts. *Proc Natl Acad Sci U S A* **94**: 7388-7393.
- de Laat, W. L., Jaspers, N. G., and Hoeijmakers, J. H. (1999). Molecular mechanism of nucleotide excision repair. *Genes Dev* **13**: 768-785.
- De Rosa, M., Fasano, C., Panariello, L., Scarano, M. I., Belli, G., Iannelli, A., Ciciliano, F., and Izzo, P. (2000). Evidence for a recessive inheritance of Turcot's syndrome caused by compound heterozygous mutations within the PMS2 gene. *Oncogene* **19**: 1719-1723.
- de Wind, N., Dekker, M., Claij, N., Jansen, L., van Klink, Y., Radman, M., Riggins, G., van der Valk, M., van't Wout, K., and te Riele, H. (1999). HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. *Nat Genet* **23**: 359-362.
- Deng, G., Chen, A., Hong, J., Chae, H. S., and Kim, Y. S. (1999). Methylation of CpG in a small region of the hMLH1 promoter invariably correlates with the absence of gene expression. *Cancer Res* **59**: 2029-2033.
- Deplanque, G., Ceraline, J., Mah-Becherel, M. C., Cazenave, J. P., Bergerat, J. P., and Klein-Soyer, C. (2001). Caffeine and the G2/M block override: a concept resulting from a misleading cell kinetic delay, independent of functional p53. *Int J Cancer* **94**: 363-369.
- Deplanque, G., Vincent, F., Mah-Becherel, M. C., Cazenave, J. P., Bergerat, J. P., and Klein-Soyer, C. (2000). Caffeine does not cause override of the G2/M block induced by UVc or gamma radiation in normal human skin fibroblasts. *Br J Cancer* **83**: 346-353.

- Desjardins, P., de Muys, J. M., and Morais, R. (1986). An established avian fibroblast cell line without mitochondrial DNA. *Somat Cell Mol Genet* **12**: 133-139.
- Dietrich, C., Wallenfang, K., Oesch, F., and Wieser, R. (1997). Differences in the mechanisms of growth control in contact-inhibited and serum-deprived human fibroblasts. *Oncogene* **15**: 2743-2747.
- DiFiglia, M., Sapp, E., Chase, K. O., Davies, S. W., Bates, G. P., Vonsattel, J. P., and Aronin, N. (1997). Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* **277**: 1990-1993.
- Doroshov, J. H. (1981). Mitomycin C-enhanced superoxide and hydrogen peroxide formation in rat heart. *J Pharmacol Exp Ther* **218**: 206-211.
- Dotto, G. P. (2000). p21(WAF1/Cip1): more than a break to the cell cycle? *Biochim Biophys Acta* **1471**: M43-56.
- Dragatsis, I., Levine, M. S., and Zeitlin, S. (2000). Inactivation of Hdh in the brain and testis results in progressive neurodegeneration and sterility in mice. *Nat Genet* **26**: 300-306.
- Dronkert, M. L., and Kanaar, R. (2001). Repair of DNA interstrand cross-links. *Mutat Res* **486**: 217-247.
- Drummond, J. T., and Bellacosa, A. (2001). Human DNA mismatch repair in vitro operates independently of methylation status at CpG sites. *Nucleic Acids Res* **29**: 2234-2243.
- Drummond, J. T., Genschel, J., Wolf, E., and Modrich, P. (1997). DHFR/MSH3 amplification in methotrexate-resistant cells alters the hMutSalpha/hMutSbeta ratio and reduces the efficiency of base-base mismatch repair. *Proc Natl Acad Sci U S A* **94**: 10144-10149.
- DuBois, R. N., and Smalley, W. E. (1996). Cyclooxygenase, NSAIDs, and colorectal cancer. *J Gastroenterol* **31**: 898-906.
- Dubrova, Y. E., Plumb, M., Gutierrez, B., Boulton, E., and Jeffreys, A. J. (2000). Transgenerational mutation by radiation. *Nature* **405**: 37.
- Duguet, M. (1997). When helicase and topoisomerase meet! *J Cell Sci* **110**: 1345-1350.
- Dunkern, T. R., Fritz, G., and Kaina, B. (2001). Ultraviolet light-induced DNA damage triggers apoptosis in nucleotide excision repair-deficient cells via Bcl-2 decline and caspase-3/-8 activation. *Oncogene* **20**: 6026-6038.
- Duranteau, J., Chandel, N. S., Kulisz, A., Shao, Z., and Schumacker, P. T. (1998). Intracellular signaling by reactive oxygen species during hypoxia in cardiomyocytes. *J Biol Chem* **273**: 11619-11624.
- Duthie, S. J., and Collins, A. R. (1997). The influence of cell growth, detoxifying enzymes and DNA repair on hydrogen peroxide-mediated DNA damage (measured using the comet assay) in human cells. *Free Radic Biol Med* **22**: 717-724.
- Dutreix, M. (1997). (GT)<sub>n</sub> repetitive tracts affect several stages of RecA-promoted recombination. *J Mol Biol* **273**: 105-113.
- Eckardt-Schupp, F., and Klaus, C. (1999). Radiation inducible DNA repair processes in eukaryotes. *Biochimie* **81**: 161-171.
- Edelmann, W., Cohen, P. E., Kane, M., Lau, K., Morrow, B., Bennett, S., Umar, A., Kunkel, T., Cattoretti, G., Chaganti, R., *et al.* (1996). Meiotic pachytene arrest in MLH1-deficient mice. *Cell* **85**: 1125-1134.
- Edelmann, W., Umar, A., Yang, K., Heyer, J., Kucherlapati, M., Lia, M., Kneitz, B., Avdievich, E., Fan, K., Wong, E., *et al.* (2000). The DNA mismatch repair genes Msh3 and Msh6 cooperate in intestinal tumor suppression. *Cancer Res* **60**: 803-807.
- Edelmann, W., Yang, K., Umar, A., Heyer, J., Lau, K., Fan, K., Liedtke, W., Cohen, P. E., Kane, M. F., Lipford, J. R., *et al.* (1997). Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. *Cell* **91**: 467-477.
- Egan, K. M., Stampfer, M. J., Giovannucci, E., Rosner, B. A., and Colditz, G. A. (1996). Prospective study of regular aspirin use and the risk of breast cancer. *J Natl Cancer Inst* **88**: 988-993.
- 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 in the FMR1 gene. *Nat Genet* **8**: 88-94.
- Eichler, E. E., Macpherson, J. N., Murray, A., Jacobs, P. A., Chakravarti, A., and Nelson, D. L. (1996). Haplotype and interspersed analysis of the FMR1 CGG repeat identifies two different mutational pathways for the origin of the fragile X syndrome. *Hum Mol Genet* **5**: 319-330.
- Elledge, S. J. (1996). Cell cycle checkpoints: preventing an identity crisis. *Science* **274**: 1664-1672.

- Ellison, A. R., Lofing, J., and Bitter, G. A. (2001). Functional analysis of human MLH1 and MSH2 missense variants and hybrid human-yeast MLH1 proteins in *Saccharomyces cerevisiae*. *Hum Mol Genet* **10**: 1889-1900.
- Enright, H. U., Miller, W. J., and Hebbel, R. P. (1992). Nucleosomal histone protein protects DNA from iron-mediated damage. *Nucleic Acids Res* **20**: 3341-3346.
- Epe, B. (1996). DNA damage profiles induced by oxidizing agents. *Rev Physiol Biochem Pharmacol* **127**: 223-249.
- Eriksson, M., Ansved, T., Anvret, M., and Carey, N. (2001). A mammalian radial spokehead-like gene, RSHL1, at the myotonic dystrophy-1 locus. *Biochem Biophys Res Commun* **281**: 835-841.
- Eriksson, M., Ansved, T., Edstrom, L., Anvret, M., and Carey, N. (1999). Simultaneous analysis of expression of the three myotonic dystrophy locus genes in adult skeletal muscle samples: the CTG expansion correlates inversely with DMPK and 59 expression levels, but not DMAHP levels. *Hum Mol Genet* **8**: 1053-1060.
- Fardaei, M., Larkin, K., Brook, J. D., and Hamshere, M. G. (2001). In vivo co-localisation of MBNL protein with DMPK expanded-repeat transcripts. *Nucleic Acids Res* **29**: 2766-2771.
- Fardaei, M., Rogers, M. T., Thorpe, H. M., Larkin, K., Hamshere, M. G., Harper, P. S., and Brook, J. D. (2002). Three proteins, MBNL, MBLL and MBXL, co-localize in vivo with nuclear foci of expanded-repeat transcripts in DM1 and DM2 cells. *Hum Mol Genet* **11**: 805-814.
- Fedier, A., Schwarz, V. A., Walt, H., Carpini, R. D., Haller, U., and Fink, D. (2001). Resistance to topoisomerase poisons due to loss of DNA mismatch repair. *Int J Cancer* **93**: 571-576.
- Feig, D. I., and Loeb, L. A. (1993). Mechanisms of mutation by oxidative DNA damage: reduced fidelity of mammalian DNA polymerase beta. *Biochemistry* **32**: 4466-4473.
- Fiore, M., and Degrassi, F. (1999). Dimethyl sulfoxide restores contact inhibition-induced growth arrest and inhibits cell density-dependent apoptosis in hamster cells. *Exp Cell Res* **251**: 102-110.
- Fojo, A. T., Ueda, K., Slamon, D. J., Poplack, D. G., Gottesman, M. M., and Pastan, I. (1987). Expression of a multidrug-resistance gene in human tumors and tissues. *Proc Natl Acad Sci U S A* **84**: 265-269.
- Fojtik, P., and Vorlickova, M. (2001). The fragile X chromosome (GCC) repeat folds into a DNA tetraplex at neutral pH. *Nucleic Acids Res* **29**: 4684-4690.
- Fortune, M. T. (2001) Developmental timing and the role of cis and trans acting modifiers on CTG repeat instability in murine models, Ph.D., University of Glasgow, Glasgow.
- Fortune, M. T., Vassilopoulos, C., Coolbaugh, M. I., Siciliano, M. J., and Monckton, D. G. (2000). Dramatic, expansion-biased, age-dependent, tissue-specific somatic mosaicism in a transgenic mouse model of triplet repeat instability. *Hum Mol Genet* **9**: 439-445.
- Fram, R. J., and Kufe, D. W. (1985). Inhibition of DNA excision repair and the repair of X-ray-induced DNA damage by cytosine arabinoside and hydroxyurea. *Pharmacol Ther* **31**: 165-176.
- Franchitto, A., Pichierri, P., Mosesso, P., and Palitti, F. (1998). Caffeine effect on the mitotic delay induced by G2 treatment with UVC or mitomycin C. *Mutagenesis* **13**: 499-505.
- Freudenreich, C. H., Kantrow, S. M., and Zakian, V. A. (1998). Expansion and length-dependent fragility of CTG repeats in yeast. *Science* **279**: 853-856.
- Freudenreich, C. H., Stavenhagen, J. B., and Zakian, V. A. (1997). Stability of a CTG/CAG trinucleotide repeat in yeast is dependent on its orientation in the genome. *Mol Cell Biol* **17**: 2090-2098.
- Froelich-Ammon, S. J., Gale, K. C., and Osheroff, N. (1994). Site-specific cleavage of a DNA hairpin by topoisomerase II. DNA secondary structure as a determinant of enzyme recognition/cleavage. *J Biol Chem* **269**: 7719-7725.
- Fry, M., and Loeb, L. A. (1999). Human werner syndrome DNA helicase unwinds tetrahelical structures of the fragile X syndrome repeat sequence d(CGG)<sub>n</sub>. *J Biol Chem* **274**: 12797-12802.
- Fu, Y. H., Friedman, D. L., Richards, S., Pearlman, J. A., Gibbs, R. A., Pizzuti, A., Ashizawa, T., Perryman, M. B., Scarlato, G., Fenwick, R. G., Jr., and et al. (1993). Decreased expression of myotonin-protein kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* **260**: 235-238.
- Fu, Y. H., Kuhl, D. P., Pizzuti, A., Pieretti, M., Sutcliffe, J. S., Richards, S., Verkerk, A. J., Holden, J. J., Fenwick, R. G., Jr., Warren, S. T., and et al. (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, R. G., Jr., King, J., Rajnarayan, S., Dunne, P. W., Dubel, J., Nasser, G. A., Ashizawa, T., de Jong, P., and et al. (1992). An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* **255**: 1256-1258.

- Funkhouser, E. M., and Sharp, G. B. (1995). Aspirin and reduced risk of esophageal carcinoma. *Cancer* **76**: 1116-1119.
- 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* **81**: 533-540.
- Gacy, A. M., Goellner, G. M., Spiro, C., Chen, X., Gupta, G., Bradbury, E. M., Dyer, R. B., Mikesell, M. J., Yao, J. Z., Johnson, A. J., *et al.* (1998). GAA instability in Friedreich's Ataxia shares a common, DNA-directed and intraallelic mechanism with other trinucleotide diseases. *Mol Cell* **1**: 583-593.
- Gacy, A. M., and McMurray, C. T. (1998). Influence of hairpins on template reannealing at trinucleotide repeat duplexes: a model for slipped DNA. *Biochemistry* **37**: 9426-9434.
- Gamper, H. B., Hou, Y. M., and Kmiec, E. B. (2000). Evidence for a four-strand exchange catalyzed by the RecA protein. *Biochemistry* **39**: 15272-15281.
- Garcia-Rodriguez, L. A., and Huerta-Alvarez, C. (2001). Reduced risk of colorectal cancer among long-term users of aspirin and nonaspirin nonsteroidal antiinflammatory drugs. *Epidemiology* **12**: 88-93.
- Gartel, A. L., and Tyner, A. L. (1999). Transcriptional regulation of the p21(WAF1/CIP1) gene. *Exp Cell Res* **246**: 280-289.
- Gasche, C., Chang, C. L., Rhees, J., Goel, A., and Boland, C. R. (2001). Oxidative stress increases frameshift mutations in human colorectal cancer cells. *Cancer Res* **61**: 7444-7448.
- Gear, A. R. (1974). Rhodamine 6G: A potent inhibitor of mitochondrial oxidative phosphorylation. *J Biol Chem* **249**: 3628-3637.
- Gez, J. (2000). FMR3 is a novel gene associated with FRAXE CpG island and transcriptionally silent in FRAXE full mutations. *J Med Genet* **37**: 782-784.
- Gez, J., Oostra, B. A., Hockey, A., Carbonell, P., Turner, G., Haan, E. A., Sutherland, G. R., and Mulley, J. C. (1997). FMR2 expression in families with FRAXE mental retardation. *Hum Mol Genet* **6**: 435-441.
- Gellibolian, R., Bacolla, A., and Wells, R. D. (1997). Triplet repeat instability and DNA topology: an expansion model based on statistical mechanics. *J Biol Chem* **272**: 16793-16797.
- Glickman, B. W., Saddi, V. A., and Curry, J. (1994). Spontaneous mutations in mammalian cells. *Mutat Res* **304**: 19-32.
- Godde, J. S., and Wolffe, A. P. (1996). Nucleosome assembly on CTG triplet repeats. *J Biol Chem* **271**: 15222-15229.
- Goellner, G. M., Tester, D., Thibodeau, S., Almqvist, E., Goldberg, Y. P., Hayden, M. R., and McMurray, C. T. (1997). Different mechanisms underlie DNA instability in Huntington disease and colorectal cancer. *Am J Hum Genet* **60**: 879-890.
- Goldberg, Y. P., Kalchman, M. A., Metzler, M., Nasir, J., Zeisler, J., Graham, R., Koide, H. B., O'Kusky, J., Sharp, A. H., Ross, C. A., *et al.* (1996). Absence of disease phenotype and intergenerational stability of the CAG repeat in transgenic mice expressing the human Huntington disease transcript. *Hum Mol Genet* **5**: 177-185.
- Goldberg, Y. P., McMurray, C. T., Zeisler, J., Almqvist, E., Sillence, D., Richards, F., Gacy, A. M., Buchanan, J., Telenius, H., and Hayden, M. R. (1995). Increased instability of intermediate alleles in families with sporadic Huntington disease compared to similar sized intermediate alleles in the general population. *Hum Mol Genet* **4**: 1911-1918.
- Gordenin, D. A., Kunkel, T. A., and Resnick, M. A. (1997). Repeat expansion-all in a flap? *Nat Genet* **16**: 116-118.
- Gossen, J. A., de Leeuw, W. J., Bakker, A. Q., and Vijg, J. (1993). DNA sequence analysis of spontaneous mutations at a LacZ transgene integrated on the mouse X chromosome. *Mutagenesis* **8**: 243-247.
- Gottesman, M. M., and Pastan, I. (1988). The multidrug transporter, a double-edged sword. *J Biol Chem* **263**: 12163-12166.
- Gourdon, G., Radvanyi, F., Lia, A. S., Duros, C., Blanche, M., Abitbol, M., Junien, C., and Hofmann-Radvanyi, H. (1997). Moderate intergenerational and somatic instability of a 55-CTG repeat in transgenic mice. *Nat Genet* **15**: 190-192.
- Graczyk, E., and Usdin, K. (2000). The GAA\*TTC triplet repeat expanded in Friedreich's ataxia impedes transcription elongation by T7 RNA polymerase in a length and supercoil dependent manner. *Nucleic Acids Res* **28**: 2815-2822.
- Gracy, R. W., Talent, J. M., Kong, Y., and Conrad, C. C. (1999). Reactive oxygen species: the unavoidable environmental insult? *Mutat Res* **428**: 17-22.

- Grankvist, K., Marklund, S. L., and Taljedal, I. B. (1981). CuZn-superoxide dismutase, Mn-superoxide dismutase, catalase and glutathione peroxidase in pancreatic islets and other tissues in the mouse. *Biochem J* **199**: 393-398.
- Grant, S. (1998). Ara-C: cellular and molecular pharmacology. *Adv Cancer Res* **72**: 197-233.
- Graveland, G. A., Williams, R. S., and DiFiglia, M. (1985). Evidence for degenerative and regenerative changes in neostriatal spiny neurons in Huntington's disease. *Science* **227**: 770-773.
- Green, H. (1993). Human genetic diseases due to codon reiteration: relationship to an evolutionary mechanism. *Cell* **74**: 955-956.
- Grewal, R. P., Cancel, G., Leeflang, E. P., Durr, A., McPeck, M. S., Draghinas, D., Yao, X., Stevanin, G., Alnot, M. O., Brice, A., and Arnheim, N. (1999). French Machado-Joseph disease patients do not exhibit gametic segregation distortion: a sperm typing analysis. *Hum Mol Genet* **8**: 1779-1784.
- Gu, Y., Desai, T., Gutierrez, P. L., and Lu, A. L. (2001). Alteration of DNA base excision repair enzymes hMYH and hOGG1 in hydrogen peroxide resistant transformed human breast cells. *Med Sci Monit* **7**: 861-868.
- Gurin, C. C., Federici, M. G., Kang, L., and Boyd, J. (1999). Causes and consequences of microsatellite instability in endometrial carcinoma. *Cancer Res* **59**: 462-466.
- Gusella, J. F., and MacDonald, M. E. (2000). Molecular genetics: unmasking polyglutamine triggers in neurodegenerative disease. *Nat Rev Neurosci* **1**: 109-115.
- Gustafson-Svard, C., Lilja, I., Hallbook, O., and Sjodahl, R. (1997). Cyclo-oxygenase and colon cancer: clues to the aspirin effect? *Ann Med* **29**: 247-252.
- Gutekunst, C. A., Li, S. H., Yi, H., Mulroy, J. S., Kuemmerle, S., Jones, R., Rye, D., Ferrante, R. J., Hersch, S. M., and Li, X. J. (1999). Nuclear and neuropil aggregates in Huntington's disease: relationship to neuropathology. *J Neurosci* **19**: 2522-2534.
- Hamshere, M. G., Newman, E. E., Alwazzan, M., Athwal, B. S., and Brook, J. D. (1997). Transcriptional abnormality in myotonic dystrophy affects DMPK but not neighboring genes. *Proc Natl Acad Sci U S A* **94**: 7394-7399.
- Han, J. W., Ahn, S. H., Park, S. H., Wang, S. Y., Bae, G. U., Seo, D. W., Kwon, H. K., Hong, S., Lee, H. Y., Lee, Y. W., and Lee, H. W. (2000). Apicidin, a histone deacetylase inhibitor, inhibits proliferation of tumor cells via induction of p21WAF1/Cip1 and gelsolin. *Cancer Res* **60**: 6068-6074.
- Hansen, R. S., Canfield, T. K., Lamb, M. M., Gartler, S. M., and Laird, C. D. (1993). Association of fragile X syndrome with delayed replication of the FMR1 gene. *Cell* **73**: 1403-1409.
- Hansen, R. S., Gartler, S. M., Scott, C. R., Chen, S. H., and Laird, C. D. (1992). Methylation analysis of CGG sites in the CpG island of the human FMR1 gene. *Hum Mol Genet* **1**: 571-578.
- Harfe, B. D., and Jinks-Robertson, S. (2000). DNA mismatch repair and genetic instability. *Annu Rev Genet* **34**: 359-399.
- Harley, H. G., Rundle, S. A., Reardon, W., Myring, J., Crow, S., Brook, J. D., Harper, P. S., and Shaw, D. J. (1992). Unstable DNA sequence in myotonic dystrophy. *Lancet* **339**: 1125-1128.
- Harper, P. S. (1998). Myotonic Dystrophy as a Trinucleotide Repeat Disorder - A Clinical Perspective. In *Genetic Instabilities and Hereditary Neurological Diseases*, R. D. Wells, and S. T. Warren, eds. (San Diego, Academic Press), pp. 115-130.
- Harper, P. S., Harley, H. G., Reardon, W., and Shaw, D. J. (1992). Anticipation in myotonic dystrophy: new light on an old problem. *Am J Hum Genet* **51**: 10-16.
- Harrington, C., and Perrino, F. W. (1995). The effects of cytosine arabinoside on RNA-primed DNA synthesis by DNA polymerase alpha-primase. *J Biol Chem* **270**: 26664-26669.
- Harris, S. E., Winchester, C. L., and Johnson, K. J. (2000). Functional analysis of the homeodomain protein SIX5. *Nucleic Acids Res* **28**: 1871-1878.
- Hashida, H., Goto, J., Suzuki, T., Jeong, S., Masuda, N., Ooie, T., Tachiiri, Y., Tsuchiya, H., and Kanazawa, I. (2001). Single cell analysis of CAG repeat in brains of dentatorubral-pallidoluysian atrophy (DRPLA). *J Neurol Sci* **190**: 87-93.
- Hassig, C. A., Tong, J. K., and Schreiber, S. L. (1997). Fiber-derived butyrate and the prevention of colon cancer. *Chem Biol* **4**: 783-789.
- Hatse, S., De Clercq, E., and Balzarini, J. (1999). Role of antimetabolites of purine and pyrimidine nucleotide metabolism in tumor cell differentiation. *Biochem Pharmacol* **58**: 539-555.
- Hayashi, J., Ohta, S., Kikuchi, A., Takemitsu, M., Goto, Y., and Nonaka, I. (1991). Introduction of disease-related mitochondrial DNA deletions into HeLa cells lacking mitochondrial DNA results in mitochondrial dysfunction. *Proc Natl Acad Sci U S A* **88**: 10614-10618.

- Hellenbroich, Y., Schwinger, E., and Zuhlke, C. (2001). Limited somatic mosaicism for Friedreich's ataxia GAA triplet repeat expansions identified by small pool PCR in blood leukocytes. *Acta Neurol Scand* **103**: 188-192.
- Hendrich, B., Abbott, C., McQueen, H., Chambers, D., Cross, S., and Bird, A. (1999). Genomic structure and chromosomal mapping of the murine and human Mbd1, Mbd2, Mbd3, and Mbd4 genes. *Mamm Genome* **10**: 906-912.
- Hendrich, B., and Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Mol Cell Biol* **18**: 6538-6547.
- Henle, E. S., and Linn, S. (1997). Formation, prevention, and repair of DNA damage by iron/hydrogen peroxide. *J Biol Chem* **272**: 19095-19098.
- Henricksen, L. A., Tom, S., Liu, Y., and Bambara, R. A. (2000). Inhibition of flap endonuclease 1 by flap secondary structure and relevance to repeat sequence expansion. *J Biol Chem* **275**: 16420-16427.
- Herman, J. G., Umar, A., Polyak, K., Graff, J. R., Ahuja, N., Issa, J. P., Markowitz, S., Willson, J. K., Hamilton, S. R., Kinzler, K. W., *et al.* (1998). Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma. *Proc Natl Acad Sci U S A* **95**: 6870-6875.
- Herrmann, H., and Aebi, U. (2000). Intermediate filaments and their associates: multi-talented structural elements specifying cytoarchitecture and cytodynamics. *Curr Opin Cell Biol* **12**: 79-90.
- Heyer, J., Yang, K., Lipkin, M., Edelmann, W., and Kucherlapati, R. (1999). Mouse models for colorectal cancer. *Oncogene* **18**: 5325-5333.
- Hickman, M. J., and Samson, L. D. (1999). Role of DNA mismatch repair and p53 in signaling induction of apoptosis by alkylating agents. *Proc Natl Acad Sci U S A* **96**: 10764-10769.
- Ho, T. C., Del Priore, L. V., and Hornbeck, R. (1997). Effect of mitomycin-C on human retinal pigment epithelium in culture. *Curr Eye Res* **16**: 572-576.
- Hofmann-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 Mol Genet* **2**: 1263-1266.
- Holmberg, M., Duyckaerts, C., Durr, A., Cancel, G., Gourfinkel-An, I., Damier, P., Faucheux, B., Trottier, Y., Hirsch, E. C., Agid, Y., and Brice, A. (1998). Spinocerebellar ataxia type 7 (SCA7): a neurodegenerative disorder with neuronal intranuclear inclusions. *Hum Mol Genet* **7**: 913-918.
- Holmes, S. E., O'Hearn, E. E., McInnis, M. G., Gorelick-Feldman, D. A., Kleiderlein, J. J., Callahan, C., Kwak, N. G., Ingersoll-Ashworth, R. G., Sherr, M., Sumner, A. J., *et al.* (1999). Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. *Nat Genet* **23**: 391-392.
- Hooper, D. C. (1998). Clinical applications of quinolones. *Biochim Biophys Acta* **1400**: 45-61.
- Hsu, C. S., and Li, Y. (2002). Aspirin potently inhibits oxidative DNA strand breaks: implications for cancer chemoprevention. *Biochem Biophys Res Commun* **293**: 705-709.
- Hunter, A., Tsilfidis, C., Mettler, G., Jacob, P., Mahadevan, M., Surh, L., and Korneluk, R. (1992). The correlation of age of onset with CTG trinucleotide repeat amplification in myotonic dystrophy. *J Med Genet* **29**: 774-779.
- Iannicola, C., Moreno, S., Oliverio, S., Nardacci, R., Ciofi-Luzzatto, A., and Piacentini, M. (2000). Early alterations in gene expression and cell morphology in a mouse model of Huntington's disease. *J Neurochem* **75**: 830-839.
- Ibrahim, W., Lee, U. S., Yen, H. C., St Clair, D. K., and Chow, C. K. (2000). Antioxidant and oxidative status in tissues of manganese superoxide dismutase transgenic mice. *Free Radic Biol Med* **28**: 397-402.
- Igarashi, S., Takiyama, Y., Cancel, G., Rogaeva, E. A., Sasaki, H., Wakisaka, A., Zhou, Y. X., Takano, H., Endo, K., Sanpei, K., *et al.* (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 mechanisms of the instability of the CAG repeat. *Hum Mol Genet* **5**: 923-932.
- Iizaka, M., Furukawa, Y., Tsunoda, T., Akashi, H., Ogawa, M., and Nakamura, Y. (2002). Expression profile analysis of colon cancer cells in response to sulindac or aspirin. *Biochem Biophys Res Commun* **292**: 498-512.
- Ikeda, H., Yamaguchi, M., Sugai, S., Aze, Y., Narumiya, S., and Kakizuka, A. (1996). Expanded polyglutamine in the Machado-Joseph disease protein induces cell death in vitro and in vivo. *Nat Genet* **13**: 196-202.
- Ikeuchi, T., Sanpei, K., Takano, H., Sasaki, H., Tashiro, K., Cancel, G., Brice, A., Bird, T. D., Schellenberg, G. D., Pericak-Vance, M. A., *et al.* (1998). A novel long and unstable CAG/CTG trinucleotide repeat on chromosome 17q. *Genomics* **49**: 321-326.

- Imbert, G., Kretz, C., Johnson, K., and Mandel, J. L. (1993). Origin of the expansion mutation in myotonic dystrophy. *Nat Genet* **4**: 72-76.
- Inoue, K., Ito, S., Takai, D., Soejima, A., Shisa, H., LePecq, J. B., Segal-Bendirdjian, E., Kagawa, Y., and Hayashi, J. I. (1997a). Isolation of mitochondrial DNA-less mouse cell lines and their application for trapping mouse synaptosomal mitochondrial DNA with deletion mutations. *J Biol Chem* **272**: 15510-15515.
- Inoue, K., Takai, D., Hosaka, H., Ito, S., Shitara, H., Isobe, K., LePecq, J. B., Segal-Bendirdjian, E., and Hayashi, J. (1997b). Isolation and characterization of mitochondrial DNA-less lines from various mammalian cell lines by application of an anticancer drug, ditercalinium. *Biochem Biophys Res Commun* **239**: 257-260.
- Iwasaki, H., Huang, P., Keating, M. J., and Plunkett, W. (1997). Differential incorporation of ara-C, gemcitabine, and fludarabine into replicating and repairing DNA in proliferating human leukemia cells. *Blood* **90**: 270-278.
- Iyer, R. R., Pluciennik, A., Rosche, W. A., Sinden, R. R., and Wells, R. D. (2000). DNA polymerase III proofreading mutants enhance the expansion and deletion of triplet repeat sequences in *Escherichia coli*. *J Biol Chem* **275**: 2174-2184.
- Iyer, R. R., and Wells, R. D. (1999). Expansion and deletion of triplet repeat sequences in *Escherichia coli* occur on the leading strand of DNA replication. *J Biol Chem* **274**: 3865-3877.
- Iyer, V. N., and Szybalski, W. (1963). A molecular mechanism of mitomycin action: linking of complementary DNA strands. *Proc Natl Acad Sci U S A* **50**: 355-362.
- Jablonska, E., Goitein, R., Marcus, M., and Cedar, H. (1985). DNA hypomethylation causes an increase in DNase-I sensitivity and an advance in the time of replication of the entire inactive X chromosome. *Chromosoma* **93**: 152-156.
- Jackson, A. L., Chen, R., and Loeb, L. A. (1998). Induction of microsatellite instability by oxidative DNA damage. *Proceedings of the National Academy of Sciences of the USA* **95**: 12468-12472.
- Jackson, A. L., and Loeb, L. A. (2000). Microsatellite instability induced by hydrogen peroxide in *Escherichia coli*. *Mutat Res* **447**: 187-198.
- Jackson, S. P. (2002). Sensing and repairing DNA double-strand breaks. *Carcinogenesis* **23**: 687-696.
- Jackson-Grusby, L., Laird, P. W., Magge, S. N., Moeller, B. J., and Jaenisch, R. (1997). Mutagenicity of 5-aza-2'-deoxycytidine is mediated by the mammalian DNA methyltransferase. *Proc Natl Acad Sci U S A* **94**: 4681-4685.
- Jakubowski, W., and Bartosz, G. (2000). 2,7-dichlorofluorescein oxidation and reactive oxygen species: what does it measure? *Cell Biol Int* **24**: 757-760.
- Jakupciak, J. P., and Wells, R. D. (1999). Genetic instabilities in (CTG.CAG) repeats occur by recombination. *J Biol Chem* **274**: 23468-23479.
- Jakupciak, J. P., and Wells, R. D. (2000). Gene conversion (recombination) mediates expansions of CTG[middle dot]CAG repeats. *J Biol Chem* **275**: 40003-40013.
- Jampel, H. D. (1992). Effect of brief exposure to mitomycin C on viability and proliferation of cultured human Tenon's capsule fibroblasts. *Ophthalmology* **99**: 1471-1476.
- Jankowski, C., Nasar, F., and Nag, D. K. (2000). Meiotic instability of CAG repeat tracts occurs by double-strand break repair in yeast. *Proc Natl Acad Sci U S A* **97**: 2134-2139.
- 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., de Jong, P. J., Amemiya, C., Aslanidis, C., Shaw, D. J., Harley, H. G., Brook, J. D., Fenwick, R., Korneluk, R. G., Tsilfidis, C., and et al. (1992a). Physical and genetic characterization of the distal segment of the myotonic dystrophy area on 19q. *Genomics* **13**: 509-517.
- Jansen, G., Groenen, P. J., Bachner, D., Jap, P. H., Coerwinkel, M., Oerlemans, F., van den Broek, W., Gohlsch, B., Pette, D., Plomp, J. J., et al. (1996). Abnormal myotonic dystrophy protein kinase levels produce only mild myopathy in mice. *Nat Genet* **13**: 316-324.
- Jansen, G., Mahadevan, M., Amemiya, C., Wormskamp, N., Segers, B., Hendriks, W., O'Hoy, K., Baird, S., Sabourin, L., Lennon, G., and et al. (1992b). Characterization of the myotonic dystrophy region predicts multiple protein isoform-encoding mRNAs. *Nat Genet* **1**: 261-266.
- Jansen, G., Willems, P., Coerwinkel, M., Nillesen, W., Smeets, H., Vits, L., Howeler, C., Brunner, H., and Wieringa, B. (1994). Gonosomal mosaicism in myotonic dystrophy patients: involvement of mitotic

- events in (CTG)<sub>n</sub> repeat variation and selection against extreme expansion in sperm. *Am J Hum Genet* **54**: 575-585.
- Jaworski, A., Rosche, W. A., Gellibolian, R., Kang, S., Shimizu, M., Bowater, R. P., Sinden, R. R., and Wells, R. D. (1995). Mismatch repair in *Escherichia coli* enhances instability of (CTG)<sub>n</sub> triplet repeats from human hereditary diseases. *Proc Natl Acad Sci U S A* **92**: 11019-11023.
- Jeffreys, A. J., Tamaki, K., MacLeod, A., Monckton, D. G., Neil, D. L., and Armour, J. A. (1994). Complex gene conversion events in germline mutation at human minisatellites. *Nat Genet* **6**: 136-145.
- Jin, P., and Warren, S. T. (2000). Understanding the molecular basis of fragile X syndrome. *Hum Mol Genet* **9**: 901-908.
- Jinks-Roberston, S., Greene, C., and Chen, W. (1998). Genetic instabilities in yeast. In *Genetic Instabilities and Hereditary Neurological Diseases*, R. D. Wells, and S. T. Warren, eds. (San Diego, Academic Press), pp. 467-484.
- Jinnai, K., Sugio, T., Mitani, M., Hashimoto, K., and Takahashi, K. (1999). Elongation of (CTG)<sub>n</sub> repeats in myotonic dystrophy protein kinase gene in tumors associated with myotonic dystrophy patients. *Muscle Nerve* **22**: 1271-1274.
- Jiricny, J. (2000). Mediating mismatch repair. *Nat Genet* **24**: 6-8.
- Johnston, C., Eliez, S., Dyer-Friedman, J., Hessel, D., Glaser, B., Blasey, C., Taylor, A., and Reiss, A. (2001). Neurobehavioral phenotype in carriers of the fragile X premutation. *Am J Med Genet* **103**: 314-319.
- Jones, P. A., and Taylor, S. M. (1980). Cellular differentiation, cytidine analogs and DNA methylation. *Cell* **20**: 85-93.
- Joseph, J. T., Richards, C. S., Anthony, D. C., Upton, M., Perez-Atayde, A. R., and Greenstein, P. (1997). Congenital myotonic dystrophy pathology and somatic mosaicism. *Neurology* **49**: 1457-1460.
- Juttermann, R., Li, E., and Jaenisch, R. (1994). Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. *Proc Natl Acad Sci U S A* **91**: 11797-11801.
- Kajikawa, S., Nakayama, H., Suzuki, M., Takashima, A., Murayama, O., Nishihara, M., Takahashi, M., and Doi, K. (1998). Increased expression of rat ribosomal protein L4 mRNA in 5-azacytidine-treated PC12 cells prior to apoptosis. *Biochem Biophys Res Commun* **252**: 220-224.
- Kamath-Loeb, A. S., Loeb, L. A., Johansson, E., Burgers, P. M., and Fry, M. (2001). Interactions between the Werner syndrome helicase and DNA polymerase delta specifically facilitate copying of tetraplex and hairpin structures of the d(CGCG)<sub>n</sub> trinucleotide repeat sequence. *J Biol Chem* **276**: 16439-16446.
- Kang, D., Nishida, J., Iyama, A., Nakabeppu, Y., Furuichi, M., Fujiwara, T., Sekiguchi, M., and Takeshige, K. (1995a). Intracellular localization of 8-oxo-dGTPase in human cells, with special reference to the role of the enzyme in mitochondria. *J Biol Chem* **270**: 14659-14665.
- Kang, S., Jaworski, A., Ohshima, K., and Wells, R. D. (1995b). Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in *E. coli*. *Nat Genet* **10**: 213-218.
- 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.
- Kang, S., Ohshima, K., Shimizu, M., Amirhaeri, S., and Wells, R. D. (1995c). Pausing of DNA synthesis in vitro at specific loci in CTG and CGG triplet repeats from human hereditary disease genes. *J Biol Chem* **270**: 27014-27021.
- Kang, S. G., Chung, H., Yoo, Y. D., Lee, J. G., Choi, Y. I., and Yu, Y. S. (2001). Mechanism of growth inhibitory effect of Mitomycin-C on cultured human retinal pigment epithelial cells: apoptosis and cell cycle arrest. *Curr Eye Res* **22**: 174-181.
- Kao, G. D., McKenna, W. G., and Yen, T. J. (2001). Detection of repair activity during the DNA damage-induced G2 delay in human cancer cells. *Oncogene* **20**: 3486-3496.
- Kass, S. U., Pruss, D., and Wolffe, A. P. (1997). How does DNA methylation repress transcription? *Trends Genet* **13**: 444-449.
- Kato, A., Takahashi, H., Takahashi, Y., and Matsushima, H. (1997). Inactivation of the cyclin D-dependent kinase in the rat fibroblast cell line, 3Y1, induced by contact inhibition. *J Biol Chem* **272**: 8065-8070.
- Kaufmann, S. H. (1998). Cell death induced by topoisomerase-targeted drugs: more questions than answers. *Biochim Biophys Acta* **1400**: 195-211.
- Kaufmann, W. E., Abrams, M. T., Chen, W., and Reiss, A. L. (1999). Genotype, molecular phenotype, and cognitive phenotype: correlations in fragile X syndrome. *Am J Med Genet* **83**: 286-295.

- Kaytor, M. D., Burright, E. N., Duvick, L. A., Zoghbi, H. Y., and Orr, H. T. (1997). Increased trinucleotide repeat instability with advanced maternal age. *Hum Mol Genet* **6**: 2135-2139.
- Kennedy, L., and Shelbourne, P. F. (2000). Dramatic mutation instability in HD mouse striatum: does polyglutamine load contribute to cell-specific vulnerability in Huntington's disease? *Hum Mol Genet* **9**: 2539-2544.
- Kenneson, A., Zhang, F., Hagedorn, C. H., and Warren, S. T. (2001). Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. *Hum Mol Genet* **10**: 1449-1454.
- Khajavi, M., Tari, A. M., Patel, N. B., Tsuji, K., Siwak, D. R., Meistrich, M. L., Terry, N. H., and Ashizawa, T. (2001). 'Mitotic drive' of expanded CTG repeats in myotonic dystrophy type 1 (DM1). *Hum Mol Genet* **10**: 855-863.
- Kim, J. S., Lee, S., Lee, T., Lee, Y. W., and Trepel, J. B. (2001). Transcriptional activation of p21(WAF1/CIP1) by apicidin, a novel histone deacetylase inhibitor. *Biochem Biophys Res Commun* **281**: 866-871.
- Kim, N., Bozek, G., Lo, J. C., and Storb, U. (1999). Different mismatch repair deficiencies all have the same effects on somatic hypermutation: intact primary mechanism accompanied by secondary modifications. *J Exp Med* **190**: 21-30.
- King, M. P., and Attardi, G. (1989). Human cells lacking mtDNA: repopulation with exogenous mitochondria by complementation. *Science* **246**: 500-503.
- Kinoshita, M., Igarashi, A., Komori, T., Tamura, H., Hayashi, M., Kinoshita, K., Deguchi, T., and Hirose, K. (1997). Differences in CTG triplet repeat expansions in an ovarian cancer and cyst from a patient with myotonic dystrophy. *Muscle Nerve* **20**: 622-624.
- Kizaki, H., Ohnishi, Y., Azuma, Y., Mizuno, Y., and Ohsaka, F. (1992). 1-beta-D-arabinosylcytosine and 5-azacytidine induce internucleosomal DNA fragmentation and cell death in thymocytes. *Immunopharmacology* **24**: 219-227.
- Klement, I. A., Skinner, P. J., Kaytor, M. D., Yi, H., Hersch, S. M., Clark, H. B., Zoghbi, H. Y., and Orr, H. T. (1998). Ataxin-1 nuclear localization and aggregation: role in polyglutamine-induced disease in SCA1 transgenic mice. *Cell* **95**: 41-53.
- Klesert, T. R., Cho, D. H., Clark, J. I., Maylie, J., Adelman, J., Snider, L., Yuen, E. C., Soriano, P., and Tapscott, S. J. (2000). Mice deficient in Six5 develop cataracts: implications for myotonic dystrophy. *Nat Genet* **25**: 105-109.
- Klesert, T. R., Otten, A. D., Bird, T. D., and Tapscott, S. J. (1997). Trinucleotide repeat expansion at the myotonic dystrophy locus reduces expression of DMAHP. *Nat Genet* **16**: 402-406.
- Klungland, A., Rosewell, I., Hollenbach, S., Larsen, E., Daly, G., Epe, B., Seeberg, E., Lindahl, T., and Barnes, D. E. (1999). Accumulation of premutagenic DNA lesions in mice defective in removal of oxidative base damage. *Proc Natl Acad Sci U S A* **96**: 13300-13305.
- Knight, S. J., Flannery, A. V., Hirst, M. C., Campbell, L., Christodoulou, Z., Phelps, S. R., Pointon, J., Middleton-Price, H. R., Barnicoat, A., Pembrey, M. E., and et al. (1993). Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. *Cell* **74**: 127-134.
- Koch, K. S., and Leffert, H. L. (1998). Giant hairpins formed by CUG repeats in myotonic dystrophy messenger RNAs might sterically block RNA export through nuclear pores. *J Theor Biol* **192**: 505-514.
- Koide, R., Ikeuchi, T., Onodera, O., Tanaka, H., Igarashi, S., Endo, K., Takahashi, H., Kondo, R., Ishikawa, A., Hayashi, T., and et al. (1994). Unstable expansion of CAG repeat in hereditary dentatorubral-pallidolusian atrophy (DRPLA). *Nat Genet* **6**: 9-13.
- Koide, R., Kobayashi, S., Shimohata, T., Ikeuchi, T., Maruyama, M., Saito, M., Yamada, M., Takahashi, H., and Tsuji, S. (1999). A neurological disease caused by an expanded CAG trinucleotide repeat in the TATA-binding protein gene: a new polyglutamine disease? *Hum Mol Genet* **8**: 2047-2053.
- Kono, H., Rusyn, I., Yin, M., Gabele, E., Yamashina, S., Dikalova, A., Kadiiska, M. B., Connor, H. D., Mason, R. P., Segal, B. H., et al. (2000). NADPH oxidase-derived free radicals are key oxidants in alcohol-induced liver disease. *J Clin Invest* **106**: 867-872.
- Koob, M. D., Moseley, M. L., Schut, L. J., Benzow, K. A., Bird, T. D., Day, J. W., and Ranum, L. P. (1999). An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nat Genet* **21**: 379-384.
- Kooy, R. F., Willemsen, R., and Oostra, B. A. (2000). Fragile X syndrome at the turn of the century. *Mol Med Today* **6**: 193-198.

- Korade-Mirnic, Z., Tarleton, J., Servidei, S., Casey, R. R., Gennarelli, M., Pegoraro, E., Angelini, C., and Hoffman, E. P. (1999). Myotonic dystrophy: tissue-specific effect of somatic CTG expansions on allele-specific DMAHP/SIX5 expression. *Hum Mol Genet* **8**: 1017-1023.
- Kovtun, I. V., and McMurray, C. T. (2001). Trinucleotide expansion in haploid germ cells by gap repair. *Nat Genet* **27**: 407-411.
- Kovtun, I. V., Therneau, T. M., and McMurray, C. T. (2000). Gender of the embryo contributes to CAG instability in transgenic mice containing a Huntington's disease gene. *Hum Mol Genet* **9**: 2767-2775.
- Kowaltowski, A. J., Castilho, R. F., and Vercesi, A. E. (2001). Mitochondrial permeability transition and oxidative stress. *FEBS Lett* **495**: 12-15.
- Krahe, R., Ashizawa, T., Abbruzzese, C., Roeder, E., Carango, P., Giacanelli, M., Funanage, V. L., and Siciliano, M. J. (1995). Effect of myotonic dystrophy trinucleotide repeat expansion on DMPK transcription and processing. *Genomics* **28**: 1-14.
- 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.
- Krokan, H. E., Standal, R., and Slupphaug, G. (1997). DNA glycosylases in the base excision repair of DNA. *Biochem J* **325**: 1-16.
- Kuhn-Hallek, I., Sage, D. R., Stein, L., Groelle, H., and Fingeroth, J. D. (1995). Expression of recombination activating genes (RAG-1 and RAG-2) in Epstein-Barr virus-bearing B cells. *Blood* **85**: 1289-1299.
- La Spada, A. R., Peterson, K. R., Meadows, S. A., McClain, M. E., Jeng, G., Chmelar, R. S., Haugen, H. A., Chen, K., Singer, M. J., Moore, D., *et al.* (1998). Androgen receptor YAC transgenic mice carrying CAG 45 alleles show trinucleotide repeat instability. *Hum Mol Genet* **7**: 959-967.
- 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.
- Lanson, N. A., Jr., Egeland, D. B., Royals, B. A., and Claycomb, W. C. (2000). The MRE11-NBS1-RAD50 pathway is perturbed in SV40 large T antigen-immortalized AT-1, AT-2 and HL-1 cardiomyocytes. *Nucleic Acids Res* **28**: 2882-2892.
- Larminat, F., Cambois, G., Zdzienicka, M. Z., and Defais, M. (1998). Lack of correlation between repair of DNA interstrand cross-links and hypersensitivity of hamster cells towards mitomycin C and cisplatin. *FEBS Lett* **437**: 97-100.
- Larsen, A. K., and Skladanowski, A. (1998). Cellular resistance to topoisomerase-targeted drugs: from drug uptake to cell death. *Biochim Biophys Acta* **1400**: 257-274.
- Lau, C. C., and Pardee, A. B. (1982). Mechanism by which caffeine potentiates lethality of nitrogen mustard. *Proc Natl Acad Sci U S A* **79**: 2942-2946.
- Lavedan, C., Hofmann-Radvanyi, H., Shelbourne, P., Rabes, J. P., Duros, C., Savoy, D., Dehaupas, I., Luce, S., Johnson, K., and Junien, C. (1993). Myotonic dystrophy: size- and sex-dependent dynamics of CTG meiotic instability, and somatic mosaicism. *Am J Hum Genet* **52**: 875-883.
- Leadon, S. A., and Avrutskaya, A. V. (1997). Differential involvement of the human mismatch repair proteins, hMLH1 and hMSH2, in transcription-coupled repair. *Cancer Res* **57**: 3784-3791.
- Leeflang, E. P., McPeck, M. S., and Arnheim, N. (1996). Analysis of meiotic segregation, using single-sperm typing: meiotic drive at the myotonic dystrophy locus. *Am J Hum Genet* **59**: 896-904.
- Leeflang, E. P., Tavare, S., Marjoram, P., Neal, C. O., Srinidhi, J., MacFarlane, H., MacDonald, M. E., Gusella, J. F., de Young, M., Wexler, N. S., and Arnheim, N. (1999). Analysis of germline mutation spectra at the Huntington's disease locus supports a mitotic mutation mechanism. *Hum Mol Genet* **8**: 173-183.
- Leeflang, E. P., Zhang, L., Tavare, S., Hubert, R., Srinidhi, J., MacDonald, M. E., Myers, R. H., de Young, M., Wexler, N. S., Gusella, J. F., and *et al.* (1995). Single sperm analysis of the trinucleotide repeats in the Huntington's disease gene: quantification of the mutation frequency spectrum. *Hum Mol Genet* **4**: 1519-1526.
- Lehmann, A. R. (1998). Dual functions of DNA repair genes: molecular, cellular, and clinical implications. *Bioessays* **20**: 146-155.
- LeProust, E. M., Pearson, C. E., Sinden, R. R., and Gao, X. (2000). Unexpected formation of parallel duplex in GAA and TTC trinucleotide repeats of Friedreich's ataxia. *J Mol Biol* **302**: 1063-1080.
- Leroy, C., Mann, C., and Marsolier, M. C. (2001). Silent repair accounts for cell cycle specificity in the signaling of oxidative DNA lesions. *Embo J* **20**: 2896-2906.
- Levine, A. J. (1997). p53, the cellular gatekeeper for growth and division. *Cell* **88**: 323-331.

- Lewis, R. J., Singh, O. M., Smith, C. V., Skarzynski, T., Maxwell, A., Wonacott, A. J., and Wigley, D. B. (1996). The nature of inhibition of DNA gyrase by the coumarins and the cyclothialidines revealed by X-ray crystallography. *Embo J* **15**: 1412-1420.
- Li, H., Li, S. H., Cheng, A. L., Mangiarini, L., Bates, G. P., and Li, X. J. (1999). Ultrastructural localization and progressive formation of neuropil aggregates in Huntington's disease transgenic mice. *Hum Mol Genet* **8**: 1227-1236.
- Li, H., Li, S. H., Johnston, H., Shelbourne, P. F., and Li, X. J. (2000a). Amino-terminal fragments of mutant huntingtin show selective accumulation in striatal neurons and synaptic toxicity. *Nat Genet* **25**: 385-389.
- Li, H., Li, S. H., Yu, Z. X., Shelbourne, P., and Li, X. J. (2001). Huntingtin aggregate-associated axonal degeneration is an early pathological event in Huntington's disease mice. *J Neurosci* **21**: 8473-8481.
- Li, M., Lotan, R., Levin, B., Tahara, E., Lippman, S. M., and Xu, X. C. (2000b). Aspirin induction of apoptosis in esophageal cancer: a potential for chemoprevention. *Cancer Epidemiol Biomarkers Prev* **9**: 545-549.
- Li, S. H., Lam, S., Cheng, A. L., and Li, X. J. (2000c). Intracellular huntingtin increases the expression of caspase-1 and induces apoptosis. *Hum Mol Genet* **9**: 2859-2867.
- Li, S. H., Schilling, G., Young, W. S., 3rd, Li, X. J., Margolis, R. L., Stine, O. C., Wagster, M. V., Abbott, M. H., Franz, M. L., Ranen, N. G., and et al. (1993). Huntington's disease gene (IT15) is widely expressed in human and rat tissues. *Neuron* **11**: 985-993.
- Lia, A. S., Seznec, H., Hoffmann-Radvany, H., Radvany, F., Duros, C., Saquet, C., Blanche, M., Junien, C., and Gourdon, G. (1998). Somatic instability of the CTG repeat in mice transgenic for the myotonic dystrophy region is age dependent but not correlated to the relative intertissue transcription levels and proliferative capacities. *Hum Mol Genet* **7**: 1285-1291.
- Lin, X., Antalffy, B., Kang, D., Orr, H. T., and Zoghbi, H. Y. (2000a). Polyglutamine expansion down-regulates specific neuronal genes before pathologic changes in SCA1. *Nat Neurosci* **3**: 157-163.
- Lin, X., Ramamurthi, K., Mishima, M., Kondo, A., and Howell, S. B. (2000b). p53 interacts with the DNA mismatch repair system to modulate the cytotoxicity and mutagenicity of hydrogen peroxide. *Mol Pharmacol* **58**: 1222-1229.
- Lipkin, S. M., Wang, V., Jacoby, R., Banerjee-Basu, S., Baxevanis, A. D., Lynch, H. T., Elliott, R. M., and Collins, F. S. (2000). MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. *Nat Genet* **24**: 27-35.
- Liquori, C. L., Ricker, K., Moseley, M. L., Jacobsen, J. F., Kress, W., Naylor, S. L., Day, J. W., and Ranum, L. P. (2001). Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. *Science* **293**: 864-867.
- Liu, H., Hewitt, S. R., and Hays, J. B. (2000). Antagonism of ultraviolet-light mutagenesis by the methyl-directed mismatch-repair system of *Escherichia coli*. *Genetics* **154**: 503-512.
- Lockshon, D., and Morris, D. R. (1983). Positively supercoiled plasmid DNA is produced by treatment of *Escherichia coli* with DNA gyrase inhibitors. *Nucleic Acids Res* **11**: 2999-3017.
- Lopez-Baena, M., Mateos, S., Pinero, J., Trinidad, O., and Cortes, F. (1998). Enhanced sensitivity to topoisomerase inhibitors in synchronous CHO cells pre-treated with 5-azacytidine. *Mutat Res* **421**: 109-116.
- Lorenzetti, D., Watase, K., Xu, B., Matzuk, M. M., Orr, H. T., and Zoghbi, H. Y. (2000). Repeat instability and motor incoordination in mice with a targeted expanded CAG repeat in the *Sca1* locus. *Hum Mol Genet* **9**: 779-785.
- Losekoot, M., Hoogendoorn, E., Olmer, R., Jansen, C. C., Oosterwijk, J. C., van den Ouweland, A. M., Halley, D. J., Warren, S. T., Willemsen, R., Oostra, B. A., and Bakker, E. (1997). Prenatal diagnosis of the fragile X syndrome: loss of mutation owing to a double recombinant or gene conversion event at the *FMR1* locus. *J Med Genet* **34**: 924-926.
- Lu, R., Nash, H. M., and Verdine, G. L. (1997). A mammalian DNA repair enzyme that excises oxidatively damaged guanines maps to a locus frequently lost in lung cancer. *Curr Biol* **7**: 397-407.
- Lu, X., Xie, W., Reed, D., Bradshaw, W. S., and Simmons, D. L. (1995). Nonsteroidal antiinflammatory drugs cause apoptosis and induce cyclooxygenases in chicken embryo fibroblasts. *Proc Natl Acad Sci U S A* **92**: 7961-7965.
- Luthi-Carter, R., Strand, A., Peters, N. L., Solano, S. M., Hollingsworth, Z. R., Menon, A. S., Frey, A. S., Spektor, B. S., Penney, E. B., Schilling, G., et al. (2000). Decreased expression of striatal signaling genes in a mouse model of Huntington's disease. *Hum Mol Genet* **9**: 1259-1271.

- Lyons-Darden, T., and Topal, M. D. (1999). Effects of temperature, Mg<sup>2+</sup> concentration and mismatches on triplet- repeat expansion during DNA replication in vitro. *Nucleic Acids Res* **27**: 2235-2240.
- Mahadevan, M., Tsilfidis, C., Sabourin, L., Shutler, G., Amemiya, C., Jansen, G., Neville, C., Narang, M., Barcelo, J., O'Hoy, K., and et al. (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. (1993). Characterization and polymerase chain reaction (PCR) detection of an Alu deletion polymorphism in total linkage disequilibrium with myotonic dystrophy. *Genomics* **15**: 446-448.
- Mangiarini, L., Sathasivam, K., Mahal, A., Mott, R., Seller, M., and Bates, G. P. (1997). Instability of highly expanded CAG repeats in mice transgenic for the Huntington's disease mutation. *Nat Genet* **15**: 197-200.
- Mangiarini, L., Sathasivam, K., Seller, M., Cozens, B., Harper, A., Hetherington, C., Lawton, M., Trotter, Y., Lehrach, H., Davies, S. W., and Bates, G. P. (1996). Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* **87**: 493-506.
- Mankodi, A., Logigian, E., Callahan, L., McClain, C., White, R., Henderson, D., Krym, M., and Thornton, C. A. (2000). Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science* **289**: 1769-1773.
- Mankodi, A., Urbinati, C. R., Yuan, Q. P., Moxley, R. T., Sansone, V., Krym, M., Henderson, D., Schalling, M., Swanson, M. S., and Thornton, C. A. (2001). Muscleblind localizes to nuclear foci of aberrant RNA in myotonic dystrophy types 1 and 2. *Hum Mol Genet* **10**: 2165-2170.
- Manley, K., Pugh, J., and Messer, A. (1999a). Instability of the CAG repeat in immortalized fibroblast cell cultures from Huntington's disease transgenic mice. *Brain Res* **835**: 74-79.
- Manley, K., Shirley, T. L., Flaherty, L., and Messer, A. (1999b). Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nat Genet* **23**: 471-473.
- Mariappan, S. V., Chen, X., Catasti, P., Bradbury, E. M., and Gupta, G. (1998). Structural studies on the unstable triplet repeats. In *Genetic Instabilities and Hereditary Neurological Diseases*, R. D. Wells, and S. T. Warren, eds. (San Diego, Academic Press), pp. 647-690.
- Mariappan, S. V., Garcoa, A. E., and Gupta, G. (1996). Structure and dynamics of the DNA hairpins formed by tandemly repeated CTG triplets associated with myotonic dystrophy. *Nucleic Acids Res* **24**: 775-783.
- Marnett, L. J. (2000). Oxyradicals and DNA damage. *Carcinogenesis* **21**: 361-370.
- Marra, G., D'Atri, S., Corti, C., Bonmassar, L., Cattaruzza, M. S., Schweizer, P., Heinemann, K., Bartosova, Z., Nystrom-Lahti, M., and Jiricny, J. (2001). Tolerance of human MSH2<sup>±</sup> lymphoblastoid cells to the methylating agent temozolomide. *Proc Natl Acad Sci U S A* **98**: 7164-7169.
- Marra, G., Iaccarino, I., Lettieri, T., Roscilli, G., Delmastro, P., and Jiricny, J. (1998). Mismatch repair deficiency associated with overexpression of the MSH3 gene. *Proc Natl Acad Sci U S A* **95**: 8568-8573.
- Marsh, J. L., Walker, H., Theisen, H., Zhu, Y. Z., Fielder, T., Purcell, J., and Thompson, L. M. (2000). Expanded polyglutamine peptides alone are intrinsically cytotoxic and cause neurodegeneration in *Drosophila*. *Hum Mol Genet* **9**: 13-25.
- Martin, E. A., Robinson, P. J., and Franklin, R. A. (2000). Oxidative stress regulates the interaction of p16 with Cdk4. *Biochem Biophys Res Commun* **275**: 764-767.
- Martorell, L., Johnson, K., Boucher, C. A., and Baiget, M. (1997). Somatic instability of the myotonic dystrophy (CTG)<sub>n</sub> repeat during human fetal development. *Hum Mol Genet* **6**: 877-880.
- Martorell, L., Martinez, J. M., Carey, N., Johnson, K., and Baiget, M. (1995). Comparison of CTG repeat length expansion and clinical progression of myotonic dystrophy over a five year period. *J Med Genet* **32**: 593-596.
- Martorell, L., Monckton, D. G., Gamez, J., and Baiget, M. (2000). Complex patterns of male germline instability and somatic mosaicism in myotonic dystrophy type 1. *Eur J Hum Genet* **8**: 423-430.
- Martorell, L., Monckton, D. G., Gamez, J., Johnson, K. J., Gich, I., de Munain, A. L., and Baiget, M. (1998). Progression of somatic CTG repeat length heterogeneity in the blood cells of myotonic dystrophy patients. *Hum Mol Genet* **7**: 307-312.
- Matilla, A., Gorbea, C., Einum, D. D., Townsend, J., Michalik, A., van Broeckhoven, C., Jensen, C. C., Murphy, K. J., Ptacek, L. J., and Fu, Y. H. (2001). Association of ataxin-7 with the proteasome subunit S4 of the 19S regulatory complex. *Hum Mol Genet* **10**: 2821-2831.
- Matsuoka, M., Tani, K., and Asano, S. (1998). Interferon-alpha-induced G1 phase arrest through up-regulated expression of CDK inhibitors, p19Ink4D and p21Cip1 in mouse macrophages. *Oncogene* **16**: 2075-2086.

- Mattern, M. R., and Scudiero, D. A. (1981). Characterization of the inhibition of replicative and repair-type DNA synthesis by novobiocin and nalidixic acid. *Biochim Biophys Acta* **653**: 248-258.
- Maurer, D. J., O'Callaghan, B. L., and Livingston, D. M. (1996). Orientation dependence of trinucleotide CAG repeat instability in *Saccharomyces cerevisiae*. *Mol Cell Biol* **16**: 6617-6622.
- McCann, J., Choi, E., Yamasaki, E., and Ames, B. N. (1975). Detection of carcinogens as mutagens in the Salmonella/microsome test: assay of 300 chemicals. *Proc Natl Acad Sci U S A* **72**: 5135-5139.
- McCullough, A. K., Dodson, M. L., and Lloyd, R. S. (1999). Initiation of base excision repair: glycosylase mechanisms and structures. *Annu Rev Biochem* **68**: 255-285.
- McGoldrick, J. P., Yeh, Y. C., Solomon, M., Essigmann, J. M., and Lu, A. L. (1995). Characterization of a mammalian homolog of the *Escherichia coli* MutY mismatch repair protein. *Mol Cell Biol* **15**: 989-996.
- McGregor, D. B., Brown, A. G., Cattanaach, P., Shepherd, W., Riach, C., Daston, D. S., and Caspary, W. J. (1989). TFT and 6TG resistance of mouse lymphoma cells to analogs of azacytidine. *Carcinogenesis* **10**: 2003-2008.
- McKay, B. C., Ljungman, M., and Rainbow, A. J. (1998). Persistent DNA damage induced by ultraviolet light inhibits p21waf1 and bax expression: implications for DNA repair, UV sensitivity and the induction of apoptosis. *Oncogene* **17**: 545-555.
- McMurray, C. T. (1999). DNA secondary structure: a common and causative factor for expansion in human disease. *Proc Natl Acad Sci U S A* **96**: 1823-1825.
- Meek, R. L., Bowman, P. D., and Daniel, C. W. (1977). Establishment of mouse embryo cells in vitro. Relationship of DNA synthesis, senescence and malignant transformation. *Exp Cell Res* **107**: 277-284.
- Mellon, I., and Champe, G. N. (1996). Products of DNA mismatch repair genes mutS and mutL are required for transcription-coupled nucleotide-excision repair of the lactose operon in *Escherichia coli*. *Proc Natl Acad Sci U S A* **93**: 1292-1297.
- Mellon, I., Rajpal, D. K., Koi, M., Boland, C. R., and Champe, G. N. (1996). Transcription-coupled repair deficiency and mutations in human mismatch repair genes. *Science* **272**: 557-560.
- Mellon, I., Spivak, G., and Hanawalt, P. C. (1987). Selective removal of transcription-blocking DNA damage from the transcribed strand of the mammalian DHFR gene. *Cell* **51**: 241-249.
- Menalled, L. B., and Chesselet, M. F. (2002). Mouse models of Huntington's disease. *Trends Pharmacol Sci* **23**: 32-39.
- Michaels, M. L., and Miller, J. H. (1992). The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine). *J Bacteriol* **174**: 6321-6325.
- Michalowski, S., Miller, J. W., Urbinati, C. R., Paliouras, M., Swanson, M. S., and Griffith, J. (1999). Visualization of double-stranded RNAs from the myotonic dystrophy protein kinase gene and interactions with CUG-binding protein. *Nucleic Acids Res* **27**: 3534-3542.
- Michel, B. (2000). Replication fork arrest and DNA recombination. *Trends Biochem Sci* **25**: 173-178.
- Mikita, T., and Beardsley, G. P. (1988). Functional consequences of the arabinosylcytosine structural lesion in DNA. *Biochemistry* **27**: 4698-4705.
- Miko, M., and Chance, B. (1975). Ethidium bromide as an uncoupler of oxidative phosphorylation. *FEBS Lett* **54**: 347-352.
- Miller, J. W., Urbinati, C. R., Teng-Umuay, P., Stenberg, M. G., Byrne, B. J., Thornton, C. A., and Swanson, M. S. (2000). Recruitment of human muscleblind proteins to (CUG)(n) expansions associated with myotonic dystrophy. *Embo J* **19**: 4439-4448.
- Miranda, S., Foncea, R., Guerrero, J., and Leighton, F. (1999). Oxidative stress and upregulation of mitochondrial biogenesis genes in mitochondrial DNA-depleted HeLa cells. *Biochem Biophys Res Commun* **258**: 44-49.
- Miret, J. J., Pessoa-Brandao, L., and Lahue, R. S. (1997). Instability of CAG and CTG trinucleotide repeats in *Saccharomyces cerevisiae*. *Mol Cell Biol* **17**: 3382-3387.
- Mitas, M. (1997). Trinucleotide repeats associated with human disease. *Nucleic Acids Res* **25**: 2245-2254.
- Miyaki, M., Nishio, J., Konishi, M., Kikuchi-Yanoshita, R., Tanaka, K., Muraoka, M., Nagato, M., Chong, J. M., Koike, M., Terada, T., *et al.* (1997). Drastic genetic instability of tumors and normal tissues in Turcotte syndrome. *Oncogene* **15**: 2877-2881.
- Modrich, P. (1997). Strand-specific mismatch repair in mammalian cells. *J Biol Chem* **272**: 24727-24730.
- Mohaghegh, P., and Hickson, I. D. (2001). DNA helicase deficiencies associated with cancer predisposition and premature ageing disorders. *Hum Mol Genet* **10**: 741-746.

- Moller, P., and Wallin, H. (1998). Adduct formation, mutagenesis and nucleotide excision repair of DNA damage produced by reactive oxygen species and lipid peroxidation product. *Mutat Res* **410**: 271-290.
- Momparler, R. L., and Bovenzi, V. (2000). DNA methylation and cancer. *J Cell Physiol* **183**: 145-154.
- Momparler, R. L., and Goodman, J. (1977). In vitro cytotoxic and biochemical effects of 5-aza-2'-deoxycytidine. *Cancer Res* **37**: 1636-1639.
- Monckton, D. G., Cayuela, M. L., Gould, F. K., Brock, G. J., Silva, R., and Ashizawa, T. (1999). Very large (CAG)<sub>n</sub> DNA repeat expansions in the sperm of two spinocerebellar ataxia type 7 males. *Hum Mol Genet* **8**: 2473-2478.
- Monckton, D. G., Coolbaugh, M. I., Ashizawa, K. T., Siciliano, M. J., and Caskey, C. T. (1997). Hypermutable myotonic dystrophy CTG repeats in transgenic mice. *Nat Genet* **15**: 193-196.
- 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.
- Montermini, L., Andermann, E., Labuda, M., Richter, A., Pandolfo, M., Cavalcanti, F., Pianese, L., Iodice, L., Farina, G., Monticelli, A., *et al.* (1997). The Friedreich ataxia GAA triplet repeat: premutation and normal alleles. *Hum Mol Genet* **6**: 1261-1266.
- Moore, H., Greenwell, P. W., Liu, C. P., Arnheim, N., and Petes, T. D. (1999). Triplet repeats form secondary structures that escape DNA repair in yeast. *Proc Natl Acad Sci U S A* **96**: 1504-1509.
- Morel, F., Debise, R., Renoux, M., Touraille, S., Ragno, M., and Alziari, S. (1999). Biochemical and molecular consequences of ethidium bromide treatment on *Drosophila* cells. *Insect Biochem Mol Biol* **29**: 835-843.
- Morel, P., Reverdy, C., Michel, B., Ehrlich, S. D., and Cassuto, E. (1998). The role of SOS and flap processing in microsatellite instability in *Escherichia coli*. *Proc Natl Acad Sci U S A* **95**: 10003-10008.
- Morrone, A., Pegoraro, E., Angelini, C., Zammarchi, E., Marconi, G., and Hoffman, E. P. (1997). RNA metabolism in myotonic dystrophy: patient muscle shows decreased insulin receptor RNA and protein consistent with abnormal insulin resistance. *J Clin Invest* **99**: 1691-1698.
- Moseley, M. L., Schut, L. J., Bird, T. D., Koob, M. D., Day, J. W., and Ranum, L. P. (2000). SCA8 CTG repeat: en masse contractions in sperm and intergenerational sequence changes may play a role in reduced penetrance. *Hum Mol Genet* **9**: 2125-2130.
- Moser, B. A., Brondello, J. M., Baber-Furnari, B., and Russell, P. (2000). Mechanism of caffeine-induced checkpoint override in fission yeast. *Mol Cell Biol* **20**: 4288-4294.
- Moses, R. E. (2001). DNA damage processing defects and disease. *Annu Rev Genomics Hum Genet* **2**: 41-68.
- Mounsey, J. P., Mistry, D. J., Ai, C. W., Reddy, S., and Moorman, J. R. (2000). Skeletal muscle sodium channel gating in mice deficient in myotonic dystrophy protein kinase. *Hum Mol Genet* **9**: 2313-2320.
- Mrak, R. E., Griffin, S. T., and Graham, D. I. (1997). Aging-associated changes in human brain. *J Neuropathol Exp Neurol* **56**: 1269-1275.
- Mu, D., Tursun, M., Duckett, D. R., Drummond, J. T., Modrich, P., and Sancar, A. (1997). Recognition and repair of compound DNA lesions (base damage and mismatch) by human mismatch repair and excision repair systems. *Mol Cell Biol* **17**: 760-769.
- Murayama, T., Takahashi, N., and Ikoma, N. (1996). Cytotoxicity and characteristics of mitomycin C. *Ophthalmic Res* **28**: 153-159.
- Mustra, D. J., Warren, A. J., and Hamilton, J. W. (2001). Preferential binding of human full-length XPA and the minimal DNA binding domain (XPA-MF122) with the mitomycin C-DNA interstrand cross-link. *Biochemistry* **40**: 7158-7164.
- Nagafuchi, S., Yanagisawa, H., Sato, K., Shirayama, T., Ohsaki, E., Bundo, M., Takeda, T., Tadokoro, K., Kondo, I., Murayama, N., and *et al.* (1994). Dentatorubral and pallidoluysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. *Nat Genet* **6**: 14-18.
- Nakamoto, M., Takebayashi, H., Kawaguchi, Y., Narumiya, S., Taniwaki, M., Nakamura, Y., Ishikawa, Y., Akiguchi, I., Kimura, J., and Kakizuka, A. (1997). A CAG/CTG expansion in the normal population. *Nat Genet* **17**: 385-386.
- Nakamura, K., Jeong, S. Y., Uchihara, T., Anno, M., Nagashima, K., Nagashima, T., Ikeda, S., Tsuji, S., and Kanazawa, I. (2001). SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum Mol Genet* **10**: 1441-1448.
- Napierala, M., and Krzyosiak, W. J. (1997). CUG repeats present in myotonin kinase RNA form metastable "slippery" hairpins. *J Biol Chem* **272**: 31079-31085.

- Napierala, M., Parniewski, P., Pluciennik, A., and Wells, R. D. (2002). Long CTG\*CAG repeat sequences markedly stimulate intramolecular recombination. *J Biol Chem* **277**: 34087-34100
- Narayanan, L., Fritzell, J. A., Baker, S. M., Liskay, R. M., and Glazer, P. M. (1997). Elevated levels of mutation in multiple tissues of mice deficient in the DNA mismatch repair gene *Pms2*. *Proc Natl Acad Sci U S A* **94**: 3122-3127.
- Nemes, J. P., Benzow, K. A., Moseley, M. L., Ranum, L. P., and Koob, M. D. (2000). The SCA8 transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (KLHL1). *Hum Mol Genet* **9**: 1543-1551.
- Neville, C. E., Mahadevan, M. S., Barcelo, 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 Mol Genet* **3**: 45-51.
- Ni, T. T., Marsischky, G. T., and Kolodner, R. D. (1999). MSH2 and MSH6 are required for removal of adenine misincorporated opposite 8-oxo-guanine in *S. cerevisiae*. *Mol Cell* **4**: 439-444.
- Nicolaides, N. C., Papadopoulos, N., Liu, B., Wei, Y. F., Carter, K. C., Ruben, S. M., Rosen, C. A., Haseltine, W. A., Fleischmann, R. D., Fraser, C. M., and et al. (1994). Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature* **371**: 75-80.
- Nishioka, K., Ohtsubo, T., Oda, H., Fujiwara, T., Kang, D., Sugimachi, K., and Nakabeppu, Y. (1999). Expression and differential intracellular localization of two major forms of human 8-oxoguanine DNA glycosylase encoded by alternatively spliced OGG1 mRNAs. *Mol Biol Cell* **10**: 1637-1652.
- Nitiss, J. L. (1998). Investigating the biological functions of DNA topoisomerases in eukaryotic cells. *Biochim Biophys Acta* **1400**: 63-81.
- Norbury, C. J., and Hickson, I. D. (2001). Cellular responses to DNA damage. *Annu Rev Pharmacol Toxicol* **41**: 367-401.
- O'Donnell, W. T., and Warren, S. T. (2002). A decade of molecular studies of fragile x syndrome. *Annu Rev Neurosci* **25**: 315-338.
- O'Hoy, K. L., Tsilfidis, C., Mahadevan, M. S., Neville, C. E., Barcelo, J., Hunter, A. G., and Korneluk, R. G. (1993). Reduction in size of the myotonic dystrophy trinucleotide repeat mutation during transmission. *Science* **259**: 809-812.
- Ogretmen, B., and Safa, A. R. (1997). Expression of the mutated p53 tumor suppressor protein and its molecular and biochemical characterization in multidrug resistant MCF-7/Adr human breast cancer cells. *Oncogene* **14**: 499-506.
- Ohno, Y., Spriggs, D., Matsukage, A., Ohno, T., and Kufe, D. (1988). Effects of 1-beta-D-arabinofuranosylcytosine incorporation on elongation of specific DNA sequences by DNA polymerase beta. *Cancer Res* **48**: 1494-1498.
- Ohshima, K., Kang, S., Larson, J. E., and Wells, R. D. (1996a). Cloning, characterization, and properties of seven triplet repeat DNA sequences. *J Biol Chem* **271**: 16773-16783.
- Ohshima, K., Kang, S., and Wells, R. D. (1996b). CTG triplet repeats from human hereditary diseases are dominant genetic expansion products in *Escherichia coli*. *J Biol Chem* **271**: 1853-1856.
- Ohshima, K., Montermini, L., Wells, R. D., and Pandolfo, M. (1998). Inhibitory effects of expanded GAA.TTC triplet repeats from intron I of the Friedreich ataxia gene on transcription and replication in vivo. *J Biol Chem* **273**: 14588-14595.
- Ohshima, K., and Wells, R. D. (1997). Hairpin formation during DNA synthesis primer realignment in vitro in triplet repeat sequences from human hereditary disease genes. *J Biol Chem* **272**: 16798-16806.
- Ona, V. O., Li, M., Vonsattel, J. P., Andrews, L. J., Khan, S. Q., Chung, W. M., Frey, A. S., Menon, A. S., Li, X. J., Stieg, P. E., et al. (1999). Inhibition of caspase-1 slows disease progression in a mouse model of Huntington's disease. *Nature* **399**: 263-267.
- Ookawara, T., Imazeki, N., Matsubara, O., Kizaki, T., Oh-Ishi, S., Nakao, C., Sato, Y., and Ohno, H. (1998). Tissue distribution of immunoreactive mouse extracellular superoxide dismutase. *Am J Physiol* **275**: C840-847.
- Ordway, J. M., Tallaksen-Greene, S., Gutekunst, C. A., Bernstein, E. M., Cearley, J. A., Wiener, H. W., Dure, L. S. t., Lindsey, R., Hersch, S. M., Jope, R. S., et al. (1997). Ectopically expressed CAG repeats cause intranuclear inclusions and a progressive late onset neurological phenotype in the mouse. *Cell* **91**: 753-763.
- Orr, H. T., Chung, M. Y., Banfi, S., Kwiatkowski, T. J., Jr., Servadio, A., Beaudet, A. L., McCall, A. E., Duvick, L. A., Ranum, L. P., and Zoghbi, H. Y. (1993). Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat Genet* **4**: 221-226.

- Otrin, V. R., McLenigan, M., Takao, M., Levine, A. S., and Protic, M. (1997). Translocation of a UV-damaged DNA binding protein into a tight association with chromatin after treatment of mammalian cells with UV light. *J Cell Sci* **110**: 1159-1168.
- Pan, X., and Leach, D. R. (2000). The roles of mutS, sbcCD and recA in the propagation of TGG repeats in *Escherichia coli*. *Nucleic Acids Res* **28**: 3178-3184.
- Paques, F., Leung, W. Y., and Haber, J. E. (1998). Expansions and contractions in a tandem repeat induced by double-strand break repair. *Mol Cell Biol* **18**: 2045-2054.
- Parker, W. B., and Cheng, Y. C. (1987). Inhibition of DNA primase by nucleoside triphosphates and their arabinofuranosyl analogs. *Mol Pharmacol* **31**: 146-151.
- Parniewski, P., Bacolla, A., Jaworski, A., and Wells, R. D. (1999). Nucleotide excision repair affects the stability of long transcribed (CTG<sup>n</sup>CAG) tracts in an orientation-dependent manner in *Escherichia coli*. *Nucleic Acids Res* **27**: 616-623.
- Parniewski, P., Jaworski, A., Wells, R. D., and Bowater, R. P. (2000). Length of CTG.CAG repeats determines the influence of mismatch repair on genetic instability. *J Mol Biol* **299**: 865-874.
- Parrow, V. C., Alestrom, P., and Gautvik, K. M. (1989). 5-azacytidine-induced alterations in the GH12C1 cells: effects on cellular morphology, chromosome structure, DNA and protein synthesis. *J Cell Sci* **93**: 533-543.
- Patel, P. I., and Isaya, G. (2001). Friedreich ataxia: from GAA triplet-repeat expansion to frataxin deficiency. *Am J Hum Genet* **69**: 15-24.
- Patrignani, P. (2000). Nonsteroidal anti-inflammatory drugs, COX-2 and colorectal cancer. *Toxicol Lett* **112-113**: 493-498.
- Paulson, H. L. (1999). Protein fate in neurodegenerative proteinopathies: polyglutamine diseases join the (mis)fold. *Am J Hum Genet* **64**: 339-345.
- Pearson, C. E., Eichler, E. E., Lorenzetti, D., Kramer, S. F., Zoghbi, H. Y., Nelson, D. L., and Sinden, R. R. (1998a). Interruptions in the triplet repeats of SCA1 and FRAXA reduce the propensity and complexity of slipped strand DNA (S-DNA) formation. *Biochemistry* **37**: 2701-2708.
- Pearson, C. E., Ewel, A., Acharya, S., Fishel, R. A., and Sinden, R. R. (1997). Human MSH2 binds to trinucleotide repeat DNA structures associated with neurodegenerative diseases. *Hum Mol Genet* **6**: 1117-1123.
- 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.
- Pearson, C. E., and Sinden, R. R. (1998a). Slipped strand DNA, dynamic mutations, and human disease. In *Genetic Instabilities and Hereditary Neurological Diseases*, R. D. Wells, and S. T. Warren, eds. (San Diego, Academic Press), pp. 585-621.
- Pearson, C. E., and Sinden, R. R. (1998b). Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA. *Curr Opin Struct Biol* **8**: 321-330.
- Pearson, C. E., Wang, Y. H., Griffith, J. D., and Sinden, R. R. (1998b). Structural analysis of slipped-strand DNA (S-DNA) formed in (CTG)<sub>n</sub>. (CAG)<sub>n</sub> repeats from the myotonic dystrophy locus. *Nucleic Acids Res* **26**: 816-823.
- Peltomaki, P. (2001a). Deficient DNA mismatch repair: a common etiologic factor for colon cancer. *Hum Mol Genet* **10**: 735-740.
- Peltomaki, P. (2001b). DNA mismatch repair and cancer. *Mutat Res* **488**: 77-85.
- Perez-Severiano, F., Rios, C., and Segovia, J. (2000). Striatal oxidative damage parallels the expression of a neurological phenotype in mice transgenic for the mutation of Huntington's disease. *Brain Res* **862**: 234-237.
- Perrino, F. W., and Mekosh, H. L. (1992). Incorporation of cytosine arabinoside monophosphate into DNA at internucleotide linkages by human DNA polymerase alpha. *J Biol Chem* **267**: 23043-23051.
- Perutz, M. F., Johnson, T., Suzuki, M., and Finch, J. T. (1994). Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc Natl Acad Sci U S A* **91**: 5355-5358.
- Petruska, J., Arnheim, N., and Goodman, M. F. (1996). Stability of intrastrand hairpin structures formed by the CAG/CTG class of DNA triplet repeats associated with neurological diseases. *Nucleic Acids Res* **24**: 1992-1998.
- Philips, A. V., Timchenko, L. T., and Cooper, T. A. (1998). Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science* **280**: 737-741.

- Pollack, V., Scheiber, K., Pfaller, W., and Schramek, H. (1997). Loss of cytokeratin expression and formation of actin stress fibers in dedifferentiated MDCK-C7 cell lines. *Biochem Biophys Res Commun* **241**: 541-547.
- Pourquier, P., Takebayashi, Y., Urasaki, Y., Gioffre, C., Kohlhagen, G., and Pommier, Y. (2000). Induction of topoisomerase I cleavage complexes by 1-beta -D- arabinofuranosylcytosine (ara-C) in vitro and in ara-C-treated cells. *Proc Natl Acad Sci U S A* **97**: 1885-1890.
- Pritsos, C. A., and Sartorelli, A. C. (1986). Generation of reactive oxygen radicals through bioactivation of mitomycin antibiotics. *Cancer Res* **46**: 3528-3532.
- Prolla, T. A., Baker, S. M., Harris, A. C., Tsao, J. L., Yao, X., Bronner, C. E., Zheng, B., Gordon, M., Reneker, J., Arnheim, N., *et al.* (1998). Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair. *Nat Genet* **18**: 276-279.
- Puccio, H., and Koenig, M. (2000). Recent advances in the molecular pathogenesis of Friedreich ataxia. *Hum Mol Genet* **9**: 887-892.
- Qin, X., Liu, L., and Gerson, S. L. (1999). Mice defective in the DNA mismatch gene PMS2 are hypersensitive to MNU induced thymic lymphoma and are partially protected by transgenic expression of human MGMT. *Oncogene* **18**: 4394-4400.
- Qin, X., Shibata, D., and Gerson, S. L. (2000). Heterozygous DNA mismatch repair gene PMS2-knockout mice are susceptible to intestinal tumor induction with N-methyl-N-nitrosourea. *Carcinogenesis* **21**: 833-838.
- Radding, C. M. (1991). Helical interactions in homologous pairing and strand exchange driven by RecA protein. *J Biol Chem* **266**: 5355-5358.
- Radicella, J. P., Dherin, C., Desmaze, C., Fox, M. S., and Boiteux, S. (1997). Cloning and characterization of hOGG1, a human homolog of the OGG1 gene of *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* **94**: 8010-8015.
- Raha, S., and Robinson, B. H. (2000). Mitochondria, oxygen free radicals, disease and ageing. *Trends Biochem Sci* **25**: 502-508.
- Razin, A. (1998). CpG methylation, chromatin structure and gene silencing-a three-way connection. *Embo J* **17**: 4905-4908.
- Reddy, S., Smith, D. B., Rich, M. M., Leferovich, J. M., Reilly, P., Davis, B. M., Tran, K., Rayburn, H., Bronson, R., Cros, D., *et al.* (1996). Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. *Nat Genet* **13**: 325-335.
- Reitmair, A. H., Schmits, R., Ewel, A., Bapat, B., Redston, M., Mitri, A., Waterhouse, P., Mittrucker, H. W., Wakeham, A., Liu, B., and *et al.* (1995). MSH2 deficient mice are viable and susceptible to lymphoid tumours. *Nat Genet* **11**: 64-70.
- Rew, D. A., and Wilson, G. D. (2000). Cell production rates in human tissues and tumours and their significance. Part 1: an introduction to the techniques of measurement and their limitations. *Eur J Surg Oncol* **26**: 227-238.
- Ricchi, P., Pignata, S., Di Popolo, A., Memoli, A., Apicella, A., Zarrilli, R., and Acquaviva, A. M. (1997). Effect of aspirin on cell proliferation and differentiation of colon adenocarcinoma Caco-2 cells. *Int J Cancer* **73**: 880-884.
- Riccio, A., Aaltonen, L. A., Godwin, A. K., Loukola, A., Percesepe, A., Salovaara, R., Masciullo, V., Genuardi, M., Paravatou-Petsotas, M., Bassi, D. E., *et al.* (1999). The DNA repair gene MBD4 (MED1) is mutated in human carcinomas with microsatellite instability. *Nat Genet* **23**: 266-268.
- Richard, G. F., Dujon, B., and Haber, J. E. (1999). Double-strand break repair can lead to high frequencies of deletions within short CAG/CTG trinucleotide repeats. *Mol Gen Genet* **261**: 871-882.
- Richard, G. F., Goellner, G. M., McMurray, C. T., and Haber, J. E. (2000). Recombination-induced CAG trinucleotide repeat expansions in yeast involve the MRE11-RAD50-XRS2 complex. *Embo J* **19**: 2381-2390.
- Richard, G. F., and Paques, F. (2000). Mini- and microsatellite expansions: the recombination connection. *EMBO Rep* **1**: 122-126.
- Richards, B., Zhang, H., Phear, G., and Meuth, M. (1997). Conditional mutator phenotypes in hMSH2-deficient tumor cell lines. *Science* **277**: 1523-1526.
- Richards, R. I. (2001). Dynamic mutations: a decade of unstable expanded repeats in human genetic disease. *Hum Mol Genet* **10**: 2187-2194.
- 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.
- Richter, M., Weiss, M., Weinberger, I., Furstenberger, G., and Marian, B. (2001). Growth inhibition and induction of apoptosis in colorectal tumor cells by cyclooxygenase inhibitors. *Carcinogenesis* **22**: 17-25.
- Ritt, M. G., Mayor, J., Wojcieszyn, J., Smith, R., Barton, C. L., and Modiano, J. F. (2000). Sustained nuclear localization of p21/WAF-1 upon growth arrest induced by contact inhibition. *Cancer Lett* **158**: 73-84.
- Roberts, R., Timchenko, N. A., Miller, J. W., Reddy, S., Caskey, C. T., Swanson, M. S., and Timchenko, L. T. (1997). Altered phosphorylation and intracellular distribution of a (CUG)<sub>n</sub> triplet repeat RNA-binding protein in patients with myotonic dystrophy and in myotonin protein kinase knockout mice. *Proc Natl Acad Sci U S A* **94**: 13221-13226.
- Rockwell, S., and Hughes, C. S. (1994). Effects of mitomycin C and porfiromycin on exponentially growing and plateau phase cultures. *Cell Prolif* **27**: 153-163.
- Roldan-Arjona, T., Wei, Y. F., Carter, K. C., Klungland, A., Anselmino, C., Wang, R. P., Augustus, M., and Lindahl, T. (1997). Molecular cloning and functional expression of a human cDNA encoding the antimutator enzyme 8-hydroxyguanine-DNA glycosylase. *Proc Natl Acad Sci U S A* **94**: 8016-8020.
- Rolfmeier, M. L., Dixon, M. J., and Lahue, R. S. (2000). Mismatch repair blocks expansions of interrupted trinucleotide repeats in yeast. *Mol Cell* **6**: 1501-1507.
- Rosenquist, T. A., Zharkov, D. O., and Grollman, A. P. (1997). Cloning and characterization of a mammalian 8-oxoguanine DNA glycosylase. *Proc Natl Acad Sci U S A* **94**: 7429-7434.
- Rossmann, T. G., and Goncharova, E. I. (1998). Spontaneous mutagenesis in mammalian cells is caused mainly by oxidative events and can be blocked by antioxidants and metallothionein. *Mutat Res* **402**: 103-110.
- Rowley, R., and Zhang, J. (1999). Caffeine-mediated override of checkpoint controls. A requirement for rhp6 (*Schizosaccharomyces pombe*). *Genetics* **152**: 61-71.
- Roza, L., van der Schans, G. P., and Lohman, P. H. (1985). The induction and repair of DNA damage and its influence on cell death in primary human fibroblasts exposed to UV-A or UV-C irradiation. *Mutat Res* **146**: 89-98.
- Rubelj, I., Huzak, M., and Brdar, B. (1999). Sudden senescence syndrome plays a major role in cell culture proliferation. *Mech Ageing Dev* **112**: 233-241.
- Ruschoff, J., Wallinger, S., Dietmaier, W., Bocker, T., Brockhoff, G., Hofstadter, F., and Fishel, R. (1998). Aspirin suppresses the mutator phenotype associated with hereditary nonpolyposis colorectal cancer by genetic selection. *Proc Natl Acad Sci U S A* **95**: 11301-11306.
- Rybak, S. L., and Murphy, R. F. (1998). Primary cell cultures from murine kidney and heart differ in endosomal pH. *J Cell Physiol* **176**: 216-222.
- Sabouri, L. A., Mahadevan, M. S., Narang, M., Lee, D. S., Surh, L. C., and Korneluk, R. G. (1993). Effect of the myotonic dystrophy (DM) mutation on mRNA levels of the DM gene. *Nat Genet* **4**: 233-238.
- Sabourin, L. A., Tamai, K., Narang, M. A., and Korneluk, R. G. (1997). Overexpression of 3'-untranslated region of the myotonic dystrophy kinase cDNA inhibits myoblast differentiation in vitro. *J Biol Chem* **272**: 29626-29635.
- Sakamoto, N., Chastain, P. D., Parniewski, P., Ohshima, K., Pandolfo, M., Griffith, J. D., and Wells, R. D. (1999). Sticky DNA: self-association properties of long GAA.TTC repeats in R.R.Y triplex structures from Friedreich's ataxia. *Mol Cell* **3**: 465-475.
- Sakumi, K., Furuichi, M., Tsuzuki, T., Kakuma, T., Kawabata, S., Maki, H., and Sekiguchi, M. (1993). Cloning and expression of cDNA for a human enzyme that hydrolyzes 8-oxo- dGTP, a mutagenic substrate for DNA synthesis. *J Biol Chem* **268**: 23524-23530.
- Samadashwily, G. M., Raca, G., and Mirkin, S. M. (1997). Trinucleotide repeats affect DNA replication in vivo. *Nat Genet* **17**: 298-304.
- Sanchez, I., Xu, C. J., Juo, P., Kakizaka, A., Blenis, J., and Yuan, J. (1999). Caspase-8 is required for cell death induced by expanded polyglutamine repeats. *Neuron* **22**: 623-633.
- Sangfelt, O., Erickson, S., Castro, J., Heiden, T., Gustafsson, A., Einhorn, S., and Grandér, D. (1999). Molecular mechanisms underlying interferon-alpha-induced G0/G1 arrest: CKI-mediated regulation of G1 Cdk-complexes and activation of pocket proteins. *Oncogene* **18**: 2798-2810.
- Sangfelt, O., Erickson, S., Einhorn, S., and Grandér, D. (1997). Induction of Cip/Kip and Ink4 cyclin dependent kinase inhibitors by interferon-alpha in hematopoietic cell lines. *Oncogene* **14**: 415-423.
- Sangfelt, O., Erickson, S., and Grandér, D. (2000). Mechanisms of interferon-induced cell cycle arrest. *Front Biosci* **5**: D479-487.

- Sano, H., Kawahito, Y., Wilder, R. L., Hashiramoto, A., Mukai, S., Asai, K., Kimura, S., Kato, H., Kondo, M., and Hla, T. (1995). Expression of cyclooxygenase-1 and -2 in human colorectal cancer. *Cancer Res* **55**: 3785-3789.
- Sanpei, K., Takano, H., Igarashi, S., Sato, T., Oyake, M., Sasaki, H., Wakisaka, A., Tashiro, K., Ishida, Y., Ikeuchi, T., *et al.* (1996). Identification of the spinocerebellar ataxia type 2 gene using a direct identification of repeat expansion and cloning technique, DIRECT. *Nat Genet* **14**: 277-284.
- Santi, D. V., Garrett, C. E., and Barr, P. J. (1983). On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs. *Cell* **33**: 9-10.
- Santi, D. V., Norment, A., and Garrett, C. E. (1984). Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytosine. *Proc Natl Acad Sci U S A* **81**: 6993-6997.
- Sargent, J. M., Elgie, A. W., Williamson, C. J., Lewandowicz, G. M., and Taylor, C. G. (2001). Circumvention of ara-C resistance by aphidicolin in blast cells from patients with AML. *Br J Cancer* **84**: 680-685.
- Sarkar, P. S., Appukuttan, B., Han, J., Ito, Y., Ai, C., Tsai, W., Chai, Y., Stout, J. T., and Reddy, S. (2000). Heterozygous loss of Six5 in mice is sufficient to cause ocular cataracts. *Nat Genet* **25**: 110-114.
- Sarkar, P. S., Chang, H. C., Boudi, F. B., and Reddy, S. (1998). CTG repeats show bimodal amplification in *E. coli*. *Cell* **95**: 531-540.
- Sato, T., Oyake, M., Nakamura, K., Nakao, K., Fukusima, Y., Onodera, O., Igarashi, S., Takano, H., Kikugawa, K., Ishida, Y., *et al.* (1999). Transgenic mice harboring a full-length human mutant DRPLA gene exhibit age-dependent intergenerational and somatic instabilities of CAG repeats comparable with those in DRPLA patients. *Hum Mol Genet* **8**: 99-106.
- Saudou, F., Finkbeiner, S., Devys, D., and Greenberg, M. E. (1998). Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. *Cell* **95**: 55-66.
- Savkur, R. S., Philips, A. V., and Cooper, T. A. (2001). Aberrant regulation of insulin receptor alternative splicing is associated with insulin resistance in myotonic dystrophy. *Nat Genet* **29**: 40-47.
- Scherer, S. J., Maier, S. M., Seifert, M., Hanselmann, R. G., Zang, K. D., Muller-Hermelink, H. K., Angel, P., Welter, C., and Schartl, M. (2000). p53 and c-Jun functionally synergize in the regulation of the DNA repair gene hMSH2 in response to UV. *J Biol Chem* **275**: 37469-37473.
- Schilling, G., Becher, M. W., Sharp, A. H., Jinnah, H. A., Duan, K., Kotzok, J. A., Slunt, H. H., Ratovitski, T., Cooper, J. K., Jenkins, N. A., *et al.* (1999). Intranuclear inclusions and neuritic aggregates in transgenic mice expressing a mutant N-terminal fragment of huntingtin. *Hum Mol Genet* **8**: 397-407.
- Schlegel, R., and Pardee, A. B. (1986). Caffeine-induced uncoupling of mitosis from the completion of DNA replication in mammalian cells. *Science* **232**: 1264-1266.
- Schmid, E., Tapscott, S., Bennett, G. S., Croop, J., Fellini, S. A., Holtzer, H., and Franke, W. W. (1979). Differential location of different types of intermediate-sized filaments in various tissues of the chicken embryo. *Differentiation* **15**: 27-40.
- Schmidt, K. H., Abbott, C. M., and Leach, D. R. (2000). Two opposing effects of mismatch repair on CTG repeat instability in *Escherichia coli*. *Mol Microbiol* **35**: 463-471.
- Schreinemachers, D. M., and Everson, R. B. (1994). Aspirin use and lung, colon, and breast cancer incidence in a prospective study. *Epidemiology* **5**: 138-146.
- Schumacher, S., Fuchs, R. P., and Bichara, M. (1998). Expansion of CTG repeats from human disease genes is dependent upon replication mechanisms in *Escherichia coli*: the effect of long patch mismatch repair revisited. *J Mol Biol* **279**: 1101-1110.
- Schumacher, S., Pinet, I., and Bichara, M. (2001). Modulation of transcription reveals a new mechanism of triplet repeat instability in *Escherichia coli*. *J Mol Biol* **307**: 39-49.
- Schweitzer, J. K., and Livingston, D. M. (1997). Destabilization of CAG trinucleotide repeat tracts by mismatch repair mutations in yeast. *Hum Mol Genet* **6**: 349-355.
- Schweitzer, J. K., and Livingston, D. M. (1999). The effect of DNA replication mutations on CAG tract stability in yeast. *Genetics* **152**: 953-963.
- Sen, S., Karthikeyan, G., and Rao, B. J. (2000). RecA realigns suboptimally paired frames of DNA repeats through a process that requires ATP hydrolysis. *Biochemistry* **39**: 10196-10206.
- Seznec, H., Agbulut, O., Sergeant, N., Savouret, C., Ghestem, A., Tabti, N., Willer, J. C., Ourth, L., Duros, C., Brisson, E., *et al.* (2001). Mice transgenic for the human myotonic dystrophy region with expanded CTG repeats display muscular and brain abnormalities. *Hum Mol Genet* **10**: 2717-2726.
- Seznec, H., Lia-Baldini, A. S., Duros, C., Fouquet, C., Lacroix, C., Hofmann-Radvanyi, H., Junien, C., and Gourdon, G. (2000). Transgenic mice carrying large human genomic sequences with expanded CTG

- repeat mimic closely the DM CTG repeat intergenerational and somatic instability. *Hum Mol Genet* **9**: 1185-1194.
- Shelbourne, P., Winqvist, R., Kunert, E., Davies, J., Leisti, J., Thiele, H., Bachmann, H., Buxton, J., Williamson, B., and Johnson, K. (1992). Unstable DNA may be responsible for the incomplete penetrance of the myotonic dystrophy phenotype. *Hum Mol Genet* **1**: 467-473.
- Shelbourne, P. F., Killeen, N., Hevner, R. F., Johnston, H. M., Tecott, L., Lewandoski, M., Ennis, M., Ramirez, L., Li, Z., Iannicola, C., *et al.* (1999). A Huntington's disease CAG expansion at the murine Hdh locus is unstable and associated with behavioural abnormalities in mice. *Hum Mol Genet* **8**: 763-774.
- Sheng, H., Shao, J., Kirkland, S. C., Isakson, P., Coffey, R. J., Morrow, J., Beauchamp, R. D., and DuBois, R. N. (1997). Inhibition of human colon cancer cell growth by selective inhibition of cyclooxygenase-2. *J Clin Invest* **99**: 2254-2259.
- Shiff, S. J., Qiao, L., Tsai, L. L., and Rigas, B. (1995). Sulindac sulfide, an aspirin-like compound, inhibits proliferation, causes cell cycle quiescence, and induces apoptosis in HT-29 colon adenocarcinoma cells. *J Clin Invest* **96**: 491-503.
- Shutler, G. G., MacKenzie, A. E., and Korneluk, R. G. (1994). The 1.5-Mb region spanning the myotonic dystrophy locus shows uniform recombination frequency. *Am J Hum Genet* **54**: 104-113.
- Sia, E. A., Dominska, M., Stefanovic, L., and Petes, T. D. (2001). Isolation and characterization of point mutations in mismatch repair genes that destabilize microsatellites in yeast. *Mol Cell Biol* **21**: 8157-8167.
- Sinden, R. R. (1999). Biological implications of the DNA structures associated with disease-causing triplet repeats. *Am J Hum Genet* **64**: 346-353.
- Sinden, R. R. (2001). Neurodegenerative diseases. Origins of instability. *Nature* **411**: 757-758.
- Sinden, R. R., Potaman, V. N., Oussatcheva, E. A., Pearson, C. E., Lyubchenko, Y. L., and Shlyakhtenko, L. S. (2002). Triplet repeat DNA structures and human genetic disease: dynamic mutations from dynamic DNA. *J Biosci* **27 Suppl 1**: 53-65.
- Singal, R., Ferris, R., Little, J. A., Wang, S. Z., and Ginder, G. D. (1997). Methylation of the minimal promoter of an embryonic globin gene silences transcription in primary erythroid cells. *Proc Natl Acad Sci U S A* **94**: 13724-13729.
- Sisodia, S. S. (1998). Nuclear inclusions in glutamine repeat disorders: are they pernicious, coincidental, or beneficial? *Cell* **95**: 1-4.
- Slupska, M. M., Baikalov, C., Luther, W. M., Chiang, J. H., Wei, Y. F., and Miller, J. H. (1996). Cloning and sequencing a human homolog (hMYH) of the Escherichia coli mutY gene whose function is required for the repair of oxidative DNA damage. *J Bacteriol* **178**: 3885-3892.
- Smits, V. A., and Medema, R. H. (2001). Checking out the G(2)/M transition. *Biochim Biophys Acta* **1519**: 1-12.
- Spector, D. L., Goldman, R. D., and Leinwand, L. A. (1998). Reduction of contamination of cultures. In *Cells: a laboratory manual* (Cold Spring Harbour, Cold Spring Harbour Laboratory Press), pp. 5.6.
- Spiro, C., Pelletier, R., Rolfsmeier, M. L., Dixon, M. J., Lahue, R. S., Gupta, G., Park, M. S., Chen, X., Mariappan, S. V., and McMurray, C. T. (1999). Inhibition of FEN-1 processing by DNA secondary structure at trinucleotide repeats. *Mol Cell* **4**: 1079-1085.
- Stark, L. A., Din, F. V., Zwacka, R. M., and Dunlop, M. G. (2001). Aspirin-induced activation of the NF-kappaB signaling pathway: a novel mechanism for aspirin-mediated apoptosis in colon cancer cells. *Faseb J* **15**: 1273-1275.
- Steffan, J. S., Bodai, L., Pallos, J., Poelman, M., McCampbell, A., Apostol, B. L., Kazantsev, A., Schmidt, E., Zhu, Y. Z., Greenwald, M., *et al.* (2001). Histone deacetylase inhibitors arrest polyglutamine-dependent neurodegeneration in Drosophila. *Nature* **413**: 739-743.
- Strong, T. V., Tagle, D. A., Valdes, J. M., Elmer, L. W., Boehm, K., Swaroop, M., Kaatz, K. W., Collins, F. S., and Albin, R. L. (1993). Widespread expression of the human and rat Huntington's disease gene in brain and nonneural tissues. *Nat Genet* **5**: 259-265.
- Suh, O., Mettlin, C., and Petrelli, N. J. (1993). Aspirin use, cancer, and polyps of the large bowel. *Cancer* **72**: 1171-1177.
- Sutcliffe, J. S., Nelson, D. L., Zhang, F., Pieretti, M., Caskey, C. T., Saxe, D., and Warren, S. T. (1992). DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum Mol Genet* **1**: 397-400.
- Sutherland, G. R., Baker, E., and Richards, R. I. (1998). Fragile sites still breaking. *Trends Genet* **14**: 501-506.
- Sutherland, G. R., and Richards, R. I. (1995). The molecular basis of fragile sites in human chromosomes. *Curr Opin Genet Dev* **5**: 323-327.

- Tachi, N., Ohya, K., Chiba, S., Nihira, H., and Minagawa, K. (1995). Muscle involvement in congenital insensitivity to pain with anhidrosis. *Pediatr Neurol* **12**: 264-266.
- Tachi, N., Ohya, K., Chiba, S., and Sato, T. (1993). Unstable DNA in a patient with a severe form of congenital myotonic dystrophy. *J Neurol Sci* **119**: 180-182.
- Takahashi, N., Murayama, T., Oda, M., and Miyakoshi, M. (1998). Cell growth inhibition and DNA incorporation of mitomycin C in cell culture. *Ophthalmic Res* **30**: 120-125.
- Takahashi-Hyodo, S. A., Sakamoto-Hojo, E. T., and Takahashi, C. S. (1999). Interaction effects of 5-azacytidine with topoisomerase II inhibitors on CHO cells, as detected by cytogenetic analysis. *Mutat Res* **431**: 13-23.
- Takano, H., Onodera, O., Takahashi, H., Igarashi, S., Yamada, M., Oyake, M., Ikeuchi, T., Koide, R., Tanaka, H., Iwabuchi, K., and Tsuji, S. (1996). Somatic mosaicism of expanded CAG repeats in brains of patients with dentatorubral-pallidoluysian atrophy: cellular population-dependent dynamics of mitotic instability. *Am J Hum Genet* **58**: 1212-1222.
- Takao, M., Aburatani, H., Kobayashi, K., and Yasui, A. (1998). Mitochondrial targeting of human DNA glycosylases for repair of oxidative DNA damage. *Nucleic Acids Res* **26**: 2917-2922.
- Takao, M., Zhang, Q. M., Yonei, S., and Yasui, A. (1999). Differential subcellular localization of human MutY homolog (hMYH) and the functional activity of adenine:8-oxoguanine DNA glycosylase. *Nucleic Acids Res* **27**: 3638-3644.
- Takiyama, Y., Sakoe, K., Amaike, M., Soutome, M., Ogawa, T., Nakano, I., and Nishizawa, M. (1999). Single sperm analysis of the CAG repeats in the gene for dentatorubral-pallidoluysian atrophy (DRPLA): the instability of the CAG repeats in the DRPLA gene is prominent among the CAG repeat diseases. *Hum Mol Genet* **8**: 453-457.
- Takiyama, Y., Sakoe, K., Soutome, M., Namekawa, M., Ogawa, T., Nakano, I., Igarashi, S., Oyake, M., Tanaka, H., Tsuji, S., and Nishizawa, M. (1997). Single sperm analysis of the CAG repeats in the gene for Machado-Joseph disease (MJD1): evidence for non-Mendelian transmission of the MJD1 gene and for the effect of the intragenic CGG/GGG polymorphism on the intergenerational instability. *Hum Mol Genet* **6**: 1063-1068.
- Tanaka, F., Reeves, M. F., Ito, Y., Matsumoto, M., Li, M., Miwa, S., Inukai, A., Yamamoto, M., Doyu, M., Yoshida, M., *et al.* (1999). Tissue-specific somatic mosaicism in spinal and bulbar muscular atrophy is dependent on CAG-repeat length and androgen receptor--gene expression level. *Am J Hum Genet* **65**: 966-973.
- 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.
- Tassone, F., Hagerman, R. J., Ikle, D. N., Dyer, P. N., Lampe, M., Willemsen, R., Oostra, B. A., and Taylor, A. K. (1999). FMRP expression as a potential prognostic indicator in fragile X syndrome. *Am J Med Genet* **84**: 250-261.
- Telenius, H., Kremer, B., Goldberg, Y. P., Theilmann, J., Andrew, S. E., Zeisler, J., Adam, S., Greenberg, C., Ives, E. J., Clarke, L. A., and *et al.* (1994). Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. *Nat Genet* **6**: 409-414.
- The Huntington's Disease Collaborative Research Group (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**: 971-983.
- Thielmann, H. W., Popanda, O., Gersbach, H., and Gilberg, F. (1993). Various inhibitors of DNA topoisomerases diminish repair-specific DNA incision in UV-irradiated human fibroblasts. *Carcinogenesis* **14**: 2341-2351.
- Thoma, F. (1999). Light and dark in chromatin repair: repair of UV-induced DNA lesions by photolyase and nucleotide excision repair. *Embo J* **18**: 6585-6598.
- Thornton, C. A., Johnson, K., and Moxley, R. T., 3rd (1994). Myotonic dystrophy patients have larger CTG expansions in skeletal muscle than in leukocytes. *Ann Neurol* **35**: 104-107.
- Thornton, C. A., Wymer, J. P., Simmons, Z., McClain, C., and Moxley, R. T., 3rd (1997). Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking DMAHP gene. *Nat Genet* **16**: 407-409.
- Thun, M. J., Namboodiri, M. M., Calle, E. E., Flanders, W. D., and Heath, C. W., Jr. (1993). Aspirin use and risk of fatal cancer. *Cancer Res* **53**: 1322-1327.
- Tian, B., White, R. J., Xia, T., Welle, S., Turner, D. H., Mathews, M. B., and Thornton, C. A. (2000). Expanded CUG repeat RNAs form hairpins that activate the double-stranded RNA-dependent protein kinase PKR. *Rna* **6**: 79-87.

- Tiefenbrun, N., Melamed, D., Levy, N., Resnitzky, D., Hoffman, I., Reed, S. I., and Kimchi, A. (1996). Alpha interferon suppresses the cyclin D3 and cdc25A genes, leading to a reversible G0-like arrest. *Mol Cell Biol* **16**: 3934-3944.
- Timchenko, L. T., and Caskey, C. T. (1999). Triplet repeat disorders: discussion of molecular mechanisms. *Cell Mol Life Sci* **55**: 1432-1447.
- Timchenko, L. T., Miller, J. W., Timchenko, N. A., DeVore, D. R., Datar, K. V., Lin, L., Roberts, R., Caskey, C. T., and Swanson, M. S. (1996a). Identification of a (CUG)<sub>n</sub> triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res* **24**: 4407-4414.
- Timchenko, L. T., Timchenko, N. A., Caskey, C. T., and Roberts, R. (1996b). Novel proteins with binding specificity for DNA CTG repeats and RNA CUG repeats: implications for myotonic dystrophy. *Hum Mol Genet* **5**: 115-121.
- Timchenko, N. A., Cai, Z. J., Welm, A. L., Reddy, S., Ashizawa, T., and Timchenko, L. T. (2001a). RNA CUG repeats sequester CUGBP1 and alter protein levels and activity of CUGBP1. *J Biol Chem* **276**: 7820-7826.
- Timchenko, N. A., Iakova, P., Cai, Z. J., Smith, J. R., and Timchenko, L. T. (2001b). Molecular basis for impaired muscle differentiation in myotonic dystrophy. *Mol Cell Biol* **21**: 6927-6938.
- Timchenko, N. A., Welm, A. L., Lu, X., and Timchenko, L. T. (1999). CUG repeat binding protein (CUGBP1) interacts with the 5' region of C/EBPbeta mRNA and regulates translation of C/EBPbeta isoforms. *Nucleic Acids Res* **27**: 4517-4525.
- Todaro, G. J., and Green, H. (1963). Quantitative studies of the mouse embryo cells in culture and their development into established lines. *J Cell Biol* **17**: 299-313.
- Toft, N. J., and Arends, M. J. (1998). DNA mismatch repair and colorectal cancer. *J Pathol* **185**: 123-129.
- Toft, N. J., Winton, D. J., Kelly, J., Howard, L. A., Dekker, M., te Riele, H., Arends, M. J., Wyllie, A. H., Margison, G. P., and Clarke, A. R. (1999). Msh2 status modulates both apoptosis and mutation frequency in the murine small intestine. *Proc Natl Acad Sci U S A* **96**: 3911-3915.
- Tomasz, M., Lipman, R., Chowdary, D., Pawlak, J., Verdine, G. L., and Nakanishi, K. (1987). Isolation and structure of a covalent cross-link adduct between mitomycin C and DNA. *Science* **235**: 1204-1208.
- Townsend, A. J., and Cheng, Y. C. (1987). Sequence-specific effects of ara-5-aza-CTP and ara-CTP on DNA synthesis by purified human DNA polymerases in vitro: visualization of chain elongation on a defined template. *Mol Pharmacol* **32**: 330-339.
- Trettel, F., Rigamonti, D., Hilditch-Maguire, P., Wheeler, V. C., Sharp, A. H., Persichetti, F., Cattaneo, E., and MacDonald, M. E. (2000). Dominant phenotypes produced by the HD mutation in STHdh(Q111) striatal cells. *Hum Mol Genet* **9**: 2799-2809.
- Tsilfidis, C., MacKenzie, A. E., Mettler, G., Barcelo, J., and Korneluk, R. G. (1992). Correlation between CTG trinucleotide repeat length and frequency of severe congenital myotonic dystrophy. *Nat Genet* **1**: 192-195.
- Tsujii, M., and DuBois, R. N. (1995). Alterations in cellular adhesion and apoptosis in epithelial cells overexpressing prostaglandin endoperoxide synthase 2. *Cell* **83**: 493-501.
- Turker, M. S., Gage, B. M., Rose, J. A., Elroy, D., Ponomareva, O. N., Stambrook, P. J., and Tischfield, J. A. (1999). A novel signature mutation for oxidative damage resembles a mutational pattern found commonly in human cancers. *Cancer Res* **59**: 1837-1839.
- Umar, A., Buermeier, A. B., Simon, J. A., Thomas, D. C., Clark, A. B., Liskay, R. M., and Kunkel, T. A. (1996). Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell* **87**: 65-73.
- Usdin, K. (1998). NGG-triplet repeats form similar intrastrand structures: implications for the triplet expansion diseases. *Nucleic Acids Res* **26**: 4078-4085.
- Usdin, K., and Grabczyk, E. (2000). DNA repeat expansions and human disease. *Cell Mol Life Sci* **57**: 914-931.
- Usdin, M. T., Shelbourne, P. F., Myers, R. M., and Madison, D. V. (1999). Impaired synaptic plasticity in mice carrying the Huntington's disease mutation. *Hum Mol Genet* **8**: 839-846.
- van Den Broek, W. J., Nelen, M. R., Wansink, D. G., Coerwinkel, M. M., te Riele, H., Groenen, P. J., and Wieringa, B. (2002). Somatic expansion behaviour of the (CTG)<sub>n</sub> repeat in myotonic dystrophy knock-in mice is differentially affected by Msh3 and Msh6 mismatch-repair proteins. *Hum Mol Genet* **11**: 191-198.

- Van Sloun, P. P., Jansen, J. G., Weeda, G., Mullenders, L. H., van Zeeland, A. A., Lohman, P. H., and Vrieling, H. (1999). The role of nucleotide excision repair in protecting embryonic stem cells from genotoxic effects of UV-induced DNA damage. *Nucleic Acids Res* **27**: 3276-3282.
- Varga-Weisz, P. D., Wilm, M., Bonte, E., Dumas, K., Mann, M., and Becker, P. B. (1997). Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* **388**: 598-602.
- Veigl, M. L., Kasturi, L., Olechnowicz, J., Ma, A. H., Lutterbaugh, J. D., Periyasamy, S., Li, G. M., Drummond, J., Modrich, P. L., Sedwick, W. D., and Markowitz, S. D. (1998). Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. *Proc Natl Acad Sci U S A* **95**: 8698-8702.
- Wang, H., Lawrence, C. W., Li, G. M., and Hays, J. B. (1999a). Specific binding of human MSH2.MSH6 mismatch-repair protein heterodimers to DNA incorporating thymine- or uracil-containing UV light photoproducts opposite mismatched bases. *J Biol Chem* **274**: 16894-16900.
- Wang, J., Pegoraro, E., Menegazzo, E., Gennarelli, M., Hoop, R. C., Angelini, C., and Hoffman, E. P. (1995). Myotonic dystrophy: evidence for a possible dominant-negative RNA mutation. *Hum Mol Genet* **4**: 599-606.
- Wang, J. C. (1996). DNA topoisomerases. *Annu Rev Biochem* **65**: 635-692.
- Wang, S. W., Norbury, C., Harris, A. L., and Toda, T. (1999b). Caffeine can override the S-M checkpoint in fission yeast. *J Cell Sci* **112**: 927-937.
- Wang, Y. H., Amirhaeri, S., Kang, S., Wells, R. D., and Griffith, J. D. (1994). Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science* **265**: 669-671.
- 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.
- Waring, M. J. (1968). Drugs which affect the structure and function of DNA. *Nature* **219**: 1320-1325.
- Warren, A. J., Ichnat, M. A., Ogdon, S. E., Rowell, E. E., and Hamilton, J. W. (1998). Binding of nuclear proteins associated with mammalian DNA repair to the mitomycin C-DNA interstrand crosslink. *Environ Mol Mutagen* **31**: 70-81.
- Weinert, T. (1998). DNA damage and checkpoint pathways: molecular anatomy and interactions with repair. *Cell* **94**: 555-558.
- Weisman-Shomer, P., Cohen, E., and Fry, M. (2000). Interruption of the fragile X syndrome expanded sequence d(CGG)(n) by interspersed d(AGG) trinucleotides diminishes the formation and stability of d(CGG)(n) tetrahelical structures. *Nucleic Acids Res* **28**: 1535-1541.
- Wells, R. D. (1996). Molecular Basis of Genetic Instability of Triplet Repeats. *J Biol Chem* **271**: 2875-2878.
- Wells, R. D., Parniewski, P., Pluciennik, A., Bacolla, A., Gellibolian, R., and Jaworski, A. (1998). Small slipped register genetic instabilities in *Escherichia coli* in triplet repeat sequences associated with hereditary neurological diseases. *J Biol Chem* **273**: 19532-19541.
- Wheeler, J. M., Beck, N. E., Kim, H. C., Tomlinson, I. P., Mortensen, N. J., and Bodmer, W. F. (1999a). Mechanisms of inactivation of mismatch repair genes in human colorectal cancer cell lines: the predominant role of hMLH1. *Proc Natl Acad Sci U S A* **96**: 10296-10301.
- Wheeler, M. D., Kono, H., Yin, M., Nakagami, M., Uesugi, T., Arteel, G. E., Gabele, E., Rusyn, I., Yamashina, S., Froh, M., *et al.* (2001). The role of Kupffer cell oxidant production in early ethanol-induced liver disease. *Free Radic Biol Med* **31**: 1544-1549.
- Wheeler, V. C., Auerbach, W., White, J. K., Srinidhi, J., Auerbach, A., Ryan, A., Duyao, M. P., Vrbanc, V., Weaver, M., Gusella, J. F., *et al.* (1999b). Length-dependent gametic CAG repeat instability in the Huntington's disease knock-in mouse. *Hum Mol Genet* **8**: 115-122.
- White, J. K., Auerbach, W., Duyao, M. P., Vonsattel, J. P., Gusella, J. F., Joyner, A. L., and MacDonald, M. E. (1997). Huntingtin is required for neurogenesis and is not impaired by the Huntington's disease CAG expansion. *Nat Genet* **17**: 404-410.
- Wiese, A. G., Pacifici, R. E., and Davies, K. J. (1995). Transient adaptation of oxidative stress in mammalian cells. *Arch Biochem Biophys* **318**: 231-240.
- Wijnen, J., van der Klift, H., Vasen, H., Khan, P. M., Menko, F., Tops, C., Meijers Heijboer, H., Lindhout, D., Moller, P., and Fodde, R. (1998). MSH2 genomic deletions are a frequent cause of HNPCC. *Nat Genet* **20**: 326-328.
- Williams, B. R. (1999). PKR; a sentinel kinase for cellular stress. *Oncogene* **18**: 6112-6120.
- Wills, P., Hickey, R., Ross, D., Cuddy, D., and Malkas, L. (1996). A novel in vitro model system for studying the action of ara-C. *Cancer Chemother Pharmacol* **38**: 366-372.

- Wills, P. W., Hickey, R., and Malkas, L. (2000). Ara-C differentially affects multiprotein forms of human cell DNA polymerase. *Cancer Chemother Pharmacol* **46**: 193-203.
- Winchester, C. L., Ferrier, R. K., Sermoni, A., Clark, B. J., and Johnson, K. J. (1999). Characterization of the expression of DMPK and SIX5 in the human eye and implications for pathogenesis in myotonic dystrophy. *Hum Mol Genet* **8**: 481-492.
- Winter, D. B., and Gearhart, P. J. (2001). Altered spectra of hypermutation in DNA repair-deficient mice. *Philos Trans R Soc Lond B Biol Sci* **356**: 5-11.
- Winter, D. B., Phung, Q. H., Umar, A., Baker, S. M., Tarone, R. E., Tanaka, K., Liskay, R. M., Kunkel, T. A., Bohr, V. A., and Gearhart, P. J. (1998). Altered spectra of hypermutation in antibodies from mice deficient for the DNA mismatch repair protein PMS2. *Proc Natl Acad Sci U S A* **95**: 6953-6958.
- Wohrle, D., Schwemmle, S., and Steinbach, P. (1996). DNA methylation and triplet repeat stability: new proposals addressing actual questions on the CGG repeat of fragile X syndrome. *Am J Med Genet* **64**: 266-267.
- Wong, A., and Cortopassi, G. (1997). mtDNA mutations confer cellular sensitivity to oxidant stress that is partially rescued by calcium depletion and cyclosporin A. *Biochem Biophys Res Commun* **239**: 139-145.
- Wong, A., Yang, J., Cavadini, P., Gellera, C., Lonnerdal, B., Taroni, F., and Cortopassi, G. (1999). The Friedreich's ataxia mutation confers cellular sensitivity to oxidant stress which is rescued by chelators of iron and calcium and inhibitors of apoptosis. *Hum Mol Genet* **8**: 425-430.
- Wong, L. J., and Ashizawa, T. (1997). Instability of the (CTG)<sub>n</sub> repeat in congenital myotonic dystrophy. *Am J Hum Genet* **61**: 1445-1448.
- Wong, L. J., 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.
- Wood, R. D. (1997). Nucleotide excision repair in mammalian cells. *J Biol Chem* **272**: 23465-23468.
- Wooster, R., Cleton-Jansen, A. M., Collins, N., Mangion, J., Cornelis, R. S., Cooper, C. S., Gusterson, B. A., Ponder, B. A., von Deimling, A., Wiestler, O. D., and et al. (1994). Instability of short tandem repeats (microsatellites) in human cancers. *Nat Genet* **6**: 152-156.
- Wörhle, 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 casual involvement of methyl-directed DNA mismatch repair in triplet repeat stability. *Hum Mol Genet* **4**: 1147-1153.
- Wu, J., Gu, L., Wang, H., Geacintov, N. E., and Li, G. M. (1999). Mismatch repair processing of carcinogen-DNA adducts triggers apoptosis. *Mol Cell Biol* **19**: 8292-8301.
- Xu, X. M., Sansores-Garcia, L., Chen, X. M., Matijevic-Aleksic, N., Du, M., and Wu, K. K. (1999). Suppression of inducible cyclooxygenase 2 gene transcription by aspirin and sodium salicylate. *Proc Natl Acad Sci U S A* **96**: 5292-5297.
- Xu, X. S., Narayanan, L., Dunklee, B., Liskay, R. M., and Glazer, P. M. (2001). Hypermutability to ionizing radiation in mismatch repair-deficient, Pms2 knockout mice. *Cancer Res* **61**: 3775-3780.
- Yakes, F. M., and Van Houten, B. (1997). Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. *Proc Natl Acad Sci U S A* **94**: 514-519.
- Yamamoto, A., Lucas, J. J., and Hen, R. (2000). Reversal of neuropathology and motor dysfunction in a conditional model of Huntington's disease. *Cell* **101**: 57-66.
- Yanagisawa, K., Kosaka, A., Iwahana, H., Nakanishi, M., and Tominaga, S. (1999). Opposite regulation of the expression of cyclin-dependent kinase inhibitors during contact inhibition. *J Biochem (Tokyo)* **125**: 36-40.
- Zaitsev, E. N., and Kowalczykowski, S. C. (1999). The simultaneous binding of two double-stranded DNA molecules by Escherichia coli RecA protein. *J Mol Biol* **287**: 21-31.
- Zeng, M., Narayanan, L., Xu, X. S., Prolla, T. A., Liskay, R. M., and Glazer, P. M. (2000). Ionizing radiation-induced apoptosis via separate Pms2- and p53- dependent pathways. *Cancer Res* **60**: 4889-4893.
- Zhang, Y., Monckton, D. G., Siciliano, M. J., Connor, T. H., and Meistrich, M. L. (2002). Age and insertion site dependence of repeat number instability of a human DM1 transgene in individual mouse sperm. *Hum Mol Genet* **11**: 791-798.
- Zheng, M., Huang, X., Smith, G. K., Yang, X., and Gao, X. (1996). Genetically unstable CXG repeats are structurally dynamic and have a high propensity for folding. An NMR and UV spectroscopic study. *J Mol Biol* **264**: 323-336.

- Zhong, N., Kajanoja, E., Smits, B., Pietrofesa, J., Curley, D., Wang, D., Ju, W., Nolin, S., Dobkin, C., Ryyanen, M., and Brown, W. T. (1996). Fragile X founder effects and new mutations in Finland. *Am J Med Genet* **64**: 226-233.
- Zhou, B. B., Chaturvedi, P., Spring, K., Scott, S. P., Johanson, R. A., Mishra, R., Mattern, M. R., Winkler, J. D., and Khanna, K. K. (2000). Caffeine abolishes the mammalian G(2)/M DNA damage checkpoint by inhibiting ataxia-telangiectasia-mutated kinase activity. *J Biol Chem* **275**: 10342-10348.
- Zhu, C., Johansson, M., and Karlsson, A. (2000). Differential incorporation of 1-beta-D-arabinofuranosylcytosine and 9- beta-D-arabinofuranosylguanine into nuclear and mitochondrial DNA. *FEBS Lett* **474**: 129-132.
- Zhuchenko, O., Bailey, J., Bonnen, P., Ashizawa, T., Stockton, D. W., Amos, C., Dobyns, W. B., Subramony, S. H., Zoghbi, H. Y., and Lee, C. C. (1997). Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the alpha 1A-voltage-dependent calcium channel. *Nat Genet* **15**: 62-69.
- Zienolddiny, S., Ryberg, D., and Haugen, A. (2000). Induction of microsatellite mutations by oxidative agents in human lung cancer cell lines. *Carcinogenesis* **21**: 1521-1526.
- Zoghbi, H. Y., and Orr, H. T. (2000). Glutamine repeats and neurodegeneration. *Annu Rev Neurosci* **23**: 217-247.