Developmental timing and the role of *cis* and *trans* acting modifiers on CTG repeat instability in murine models

by

Maria Teresa Fortune

Thesis submitted for the degree of Doctor of Philosophy

University of Glasgow

Division of Molecular Genetics

Institute of Biomedical and Life Sciences

March 2001

ProQuest Number: 13818778

ProQuest 13818778

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

M Teresa Fortune

March 2001

Dedication

To the Family Fortune

Abstract

Myotonic dystrophy type 1 (DM1), is one of a number of genetic diseases whose phenotype is associated with the expansion of a CTG•CAG repeat. In the human population the DM1 triplet repeat locus demonstrates moderate variations in repeat length, the normal range being between 5 and 37 repeats. However, repeat lengths present in patients are far larger, beyond 50 and into the thousands, with variation in a patient, within gametes and somatic mosaicism between and within tissues. The repeat instability level in DM1 patients is effected by several factors including the initial size of the progenitor allele and the patients age. Expansion over time and mutation rate variation between the somatic tissues of an individual is thought to contribute to the tissue specificity and progressive nature of the symptoms, whilst the size of the repeat inherited tends to expand through successive generations. This mechanism correlates with the incidence of 'anticipation' observed in affected families, that is the disease becomes more severe in each successive generation. To investigate the mechanism of repeat expansion murine models have been developed containing the *Dmt*-162 transgene, which consists of ~162 CTG repeats derived from the human DM1 locus. Previous studies had shown some repeat instability in the somatic tissues of young (male) mice of each line. *Dmt*-162 mice have already shown parent-of-origin-effects, with paternal repeat transmissions prone to expansion and maternal repeat transmissions prone to deletion. Further pedigree data analysis confirmed this effect and allowed the repeat transmission patterns to be analysed in more depth. The effects of the parents age on the transmitted repeat size was examined but was found not be significant in either sex.

Initially the repeat instability in a number of somatic tissues from mature mice in four of the lines generated was investigated using small pool PCR. Whilst minimal instability was observed in three of the lines, the fourth, *Dmt*-D demonstrated expansion biased instability which varied greatly between tissues. Investigation at earlier timepoints revealed the mosaicism to increase with age. Microdissection of the brain and the kidney revealed the repeat instability patterns to be region specific. Examination of the repeat mosaicism presented by the regions of the kidney observed dramatic repeat instability within the cortex. This in turn lead to the investigation of a candidate cell type for the dramatic repeat instability, the glomereli.

Investigation of transgenic transcription using reverse transcriptase PCR has identified tissue specific transcription of the repeat in three of the investigated lines and ubiquitous transcription in the fourth. The orientation of the transcribed CTG•CAG repeat has also been determined.

Finally to investigate the effect of an incomplete mismatch repair complex (MMR), *Dmt* transgenic mice were crossed on to a background deficient in the MMR protein *Pms2* and the effect on repeat transmission and somatic instability studied.

Acknowledgements

A long, long time ago in a land far, far away a man started a series of experiments, which resulted in my writing this PhD thesis, so I would like to thank Darren Monckton for being that man in the land of Texas and for being an absolute star.

Its awful having to single people out, but I'd like to thank my assessors Richard Wilson and Shireen Davies, as well as Graham Brock, Graham Hamilton, Heather Johnston, Catherine Winchester, Sarah Harris and Christine Haworth, who are or were Level 5 people and Tony McDermott and Craig Begley of the CRF and Giorgia Riboldi Tunnicliffe of the MBSU.

Now then, to thank everybody else who needs thanking and acknowledge everyone I want to acknowledge would require one hell of a long cross referenced list so what I really want to say is if you: helped me at all, at any time; encouraged me, either to get into this in the first place or just to keep going; solved my inept computer problems; established a world class lab for me to work in; made me laugh; listened; made suggestions or explained techniques; fed me, either dinner or beer; phoned just to see if it was all ok; worried, fretted and sometimes despaired on my behalf; whether or not you're a University of Wales type person or a Glasgow type person, home or London, lab or CRF. Thankyou. All. Very much.

And if you're related to me, I mean that threefold!

Table of Contents

Chapter 1: Introduction

Chapter 2: Materials and Methods

Chapter 3: Somatic instability in the *Dmt* transgenic lines

Chapter 4: Gametic instability in the *Dmt* transgenic lines

Chapter 5: Investigation of transcription levels of the *Dmt* transgenes

Chapter 6: Investigation of the stability levels of the *Dmt* transgene on a background deficient in a mismatch repair protein

Chapter 7: Discussion and future studies

References

List of tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

List of figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Abbreviations

bp	base pair
BSA	bovine serum albumin
CAG	triplet of cytosine, adenine and guanine
CCG	triplet of cytosine, cytosine and guanine
cDNA	complementary deoxyribonucleic acid
CDM	congenital myotonic dystrophy
CTG	triplet of cytosine, thymine and guanine
DEPC	diethyl pyrocarbonate
DM1	mytonic dystrophy type 1
DMPK	mytonic dystrophy protein kinase
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
DRPLA	Denatorubral pallidoluysian atrophy
DTT	dithiotheitol
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
dGTP	2'-deoxyguanosine 5'-triphosphate
dNTP	2'-deoxynucleoside 5'-triphosphate
dTTP	2'-deoxythymidine5'-triphosphate
EDTA	ethylenediaminetetracetic acid
FA	Friedreichs ataxia
6-FAM	6-carboxyfluorescein
FRAXA	fragile X syndrome
FRAXE	fragile X site E (mental retardation)
GAA	triplet of guanine, adenine and adenine
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
HNPCC	Hereditary non polyposis colon cancer
HCG	human chorionic gonadotrophin
HD	Huntington disease
IPTG	isopropylthio-β-D-galactoside
kb	kilobase

kDa	kilo Dalton
LB	Luria-Bertani
LINE	long interspersed nuclear element
Ln	natural log
M	molar
MMR	mismatch repair
mM	milli molar
MLH1	MutL homologue 1
mRNA	messenger ribonucleic acid
MSH2	MutS homologue 2
OD	optical density
OLB	oligo labelling buffer
O/N	overnight
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PMS2	post meiotic segregation gene 2
PMSG	pregnant mare serum gonadotrophin
rpm	revolutions per minute
RNA	ribonucleic acid
RNAse	ribonuclase
ROX	6-carboxy-x-rhodamine
RT-PCR	reverse transcriptase polymerase chain reaction
SBMA	spinobulbar muscular atrophy
SCA	spinocerabellar ataxia
SDS	sodium dodecyl sulphate
Six5	*sine oculis related homeobox 5*
SP-PCR	small pool polymerase chain reaction
Tris	Tris(hydroxymethyl)amino methane
USF-2	upstream stimulatory factor 2
UTR	untranslated region

UV   ultraviolet

X-gal   5-bromo-4-chloro-3-indolyl-β-D-galactoside

Chapter 1

Introduction

1.1 Triplet repeat diseases

In 1991 the mutations associated with the disorders spinal and bulbar muscular atrophy (SBMA) (La Spada *et al.*, 1991) and Fragile X syndrome (Frax A) (Fu *et al.*, 1991; Verkerk *et al.*, 1991) were identified as tracts of trinucleotide repeats that had expanded beyond a normal threshold. In the human population triplet repeat loci often show moderate variations in repeat length, but the repeat lengths present in patients are far larger with additional variation within a patient, including within gametes and somatic mosaicism within and between tissues. By 1997, when this project began, 12 gene loci of inherited genetic diseases were found to contain an expanded triplet repeat tract within patients. These include the neurodegenerative disorders, Huntington's disease (HD) (Huntington's Disease Collaborative Research Group, 1993; Zuhlke *et al.*, 1993), Denatorubral pallidoluysian atrophy (DRLPA) or Haw River Syndrome (Tsuji, 1997), and the spinocerebeller ataxias, (SCA) type 1 (Orr *et al.*, 1993), type 2 (Imbert *et al.*, 1996; Pulst *et al.*, 1996; Sanpei *et al.*, 1996), type 3 (Stevanin *et al.*, 1995), type 6 (Zhuchenko, 1997) and type 7 (David, 1997) as well as Frax E, (Flynn *et al.*, 1993; Knight *et al.*, 1993), 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) and Friedreich's ataxia (FA) (Campuzano *et al.*, 1996). The addition of spinocerebellar ataxia's (SCA) type 8 (Koob *et al.*, 1999) and type 12 (Holmes *et al.*, 1999), brings the current total number of loci associated with such inherited disorders to 14. The disorder Machado-Joseph disease (Kawaguchi *et al.*, 1994) is associated with the SCA 3 locus and there is evidence that the two disorders are distinct clinical entities (Haberhausen *et al.*, 1995).

The significance of the expanded trinucleotide mutation is the correlation of the repeat tract expanding over successive generations and the phenomenon of anticipation. Anticipation is the occurrence of an inherited disorder in which successive generations are subject to a progressively earlier age of onset (Harper *et al.*, 1992; Nagafuchi *et al.*, 1994). Until the discovery of expanded trinucleotide repeat tracts at gene loci, the anticipation observed in many inherited human diseases was thought possibly to be just the consequence of ascertainment bias (Penrose, 1948). Although Penrose, (1948), did concede that in the case of DM1, "the pedigree data required further explanation." It is now known however, that the

1

expansion of the repeat between generations correlates inversely with the reduction of the age of onset in such disorders. The size of the repeat that is transmitted tends to expand through successive generations while pedigree analysis has revealed disorder dependent sex related transmission patterns.

Although each disease is associated with an unstable triplet repeat expansion, they can themselves be subdivided into three distinct groups according to the location of the repeats within the gene (Figure 1.1 and Table 1.1):

- The repeat tracts are located within intronic regions. Such tracts are removed before the mRNA is exported to the cytoplasm.

- The repeat tract is located within an untranslated region of an exon. The expanded repeat tract may influence the processing and or export of the mRNA into the cytoplasm, but would not code for the resulting protein.

- The repeat tract is located within an exon. The repeat is translated as part of the associated protein. The translated tract diseases all contain an in frame CAG expansion and the resulting polyglutamine tracts are associated with neurodegenerative disorders.

With the exception of the polyglutamine tract associated disorders, the triplet repeat diseases appear to have little in common. They are located on a range of chromosomes and there is variety in the repeat motifs. Although the CAG motif is the most common, other untranslated expansions are transcribed as triplets of CTG, GCC, and GAA. Also not all CAG tracts are translated, the CAG expansion of SCA 12 precedes the promoter region (Holmes *et al.*, 1999), while the CAG·CTG repeat found in the 3' UTR of SCA 8 has been reported to be transcribed only as a CTG tract (Koob *et al.*, 1999).

The identification of the expansion of trinucleotide repeats as the genetic fault that underlies these disorders has lead to the investigation of expansion mechanisms.
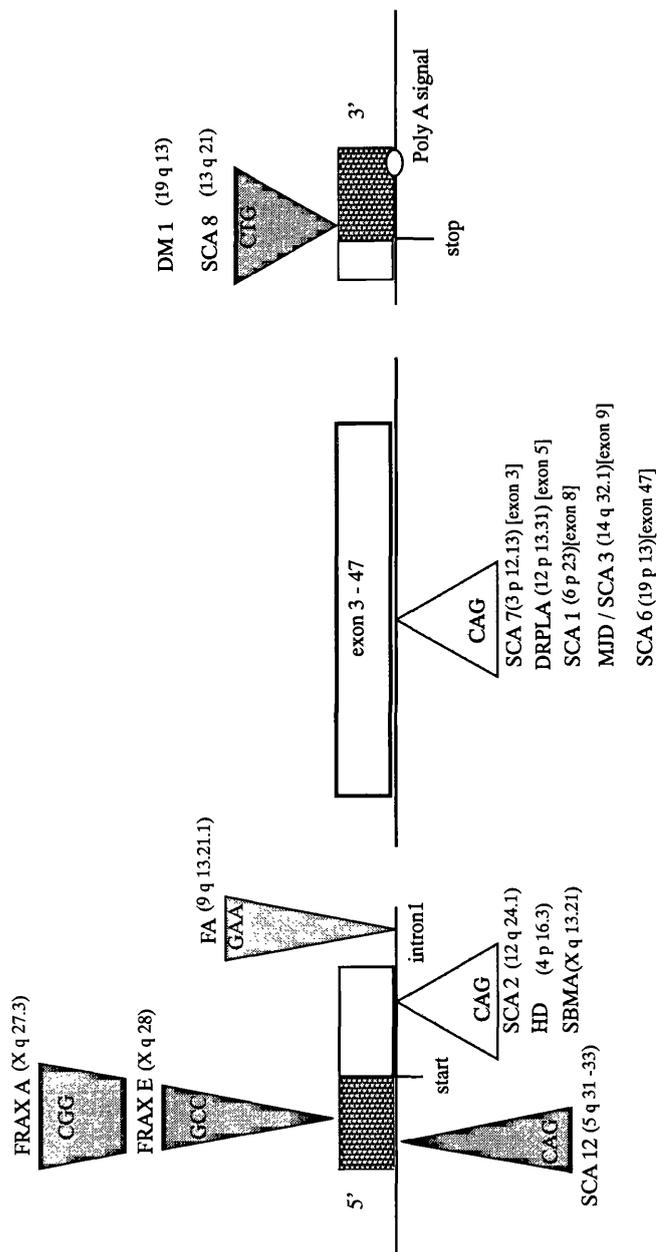
Figure 1.1 Location of the known triplet repeat diseases
Schematic diagram showing the location and sequence motif on the coding strand of each triplet repeat disorder.

Location of repeat	Disorder	Protein	Repeat	Repeat range	Reference
5' UTR					
X q 27.3	Fragile X syndrome (Frax A)	FMR-1 protein	CGG	Normal 6 - 52, Premutation 50 - 200, Mutation 230 - 1000	(Fu et al., 1991; Verkerk et al., 1991)
X q 28	Fragile XE mental retardation (Frax E)	FMR-2 protein	GCC	Normal 6 - 25, Mutation 115 - 850	(Flynn et al., 1993; Knight et al., 1993)
5 q 31-33	Spinocerebellar ataxia type 12 (SCA12)	PP2A-PR55β	CAG	Normal 7 - 12, Mutation 66 - 78	(Holmes et al., 1999)
3' UTR					
19 q 13	Myotonic dystrophy type 1 (DM1)	Myotonic dystrophy protein kinase (DMPK)	CTG	Normal 5 - 37, Mutation 50 - 3000	(Aslanidis et al., 1992; Brook et al., 1992; Buxton et al., 1992; Fu et al., 1992; Harley et al., 1992; Mahadevan et al., 1992)
13 q 21	Spinocerebellar ataxia type 8 (SCA8)		CTG	Normal 16 - 37, Mutation 107 - 127	(Koob et al., 1999)

Location of repeat	Disorder	Protein	Repeat	Repeat range		Reference
intron						
intron 1 9 q 13-21.1	Friedreichs Ataxia (FA)	Frataxin	GAA	Normal Mutation	7 - 22 200 - 900	(Campuzano et al., 1996)
exon						
exon 1 12 q 24.1	Spinocerebellar ataxia type 2 (SCA 2)	Ataxin-2	CAG	Normal Mutation	14 - 31 35 - 64	(Imbert et al., 1996; Pulst et al., 1996; Sanpei et al., 1996),
exon 1 4 p 16.3	Huntington's disease (HD)	Huntingtin	CAG	Normal Mutation	6 - 39 36 - 121	[Huntington's Disease Collaborarive Research Group, 1993 #6417; Zuhlke et al., 1993),
exon 1 X q 13 – 21	Spinal and bulbar muscular atrophy (SBMA)	Androgen receptor (AR)	CAG	Normal Mutation	11 - 33 38 - 66	(La Spada et al., 1991)
exon 3 13 p 12-13	Spinocerebellar ataxia type 7 (SCA 7)	Ataxin-7	CAG	Normal Mutation	7 - 17 38 - 130	(David, 1997)

Location of repeat	Disorder	Protein	Repeat	Repeat range		Reference
exon				Normal	Mutation	
exon 5 12 p 13.31	Denatorubalpallidoluysian atrophy (DRPLA)	Atrophin-1	CAG	6 - 35	51 - 88	(Koide et al., 1994; Nagafuchi et al., 1994)
exon 8 6 p 23	Spinocerebellar ataxia type 1 (SCA 1)	Ataxin-1	CAG	6 - 39	41 - 81	(Or et al, 1993),
exon 9 12 q 32.1	Spinocerebellar ataxia type 3 (SCA 3) / Machado Joseph Disease (MJD)	Ataxin-3	CAG	12 - 41	40 - 84	(Stevanin et al., 1995) (Kawaguchi et al., 1994)
exon 47 19 p 13	Spinocerebellar ataxia type 6 (SCA 6)	α-voltage-dependant calcium channel subunit	CAG	4 - 16	17 - 27	(Zhuchenko, 1997)

Table 1.1 Location and mutation range of triplet repeat diseases

1.2 Myotonic dystrophy (DM1)

The disease, myotonic dystrophy type 1, (DM1) was categorised in its own right in 1909, when it was defined as a clinical disorder distinct from the nonprogressive channelopathy myotonia congenita (Batten and Gibb, 1909; Steinert, 1909). The name itself describes the impaired muscle relaxation (myotonia) and muscle weakness (dystrophy) exhibited by patients. In 1989, the incidence of DM1 was reported as one in 8,000 in the Caucasian population, where it occurs most frequently (Harper, 1989). This makes it the most common adult onset muscular dystrophy, yet it is unusual among both the muscular dystrophies and the expanded triplet diseases in its multisystemic nature. It is an autosomal dominant disorder, like most but not all of the triplet repeat disorders and its genetic basis is a CTG expansion in the 3' untranslated region of the *MYOTONIC DYSTROPHY PROTEIN KINASE GENE* (*DMPK*) (Aslanidis *et al.*, 1992; Brook *et al.*, 1992; Buxton *et al.*, 1992; Fu *et al.*, 1992; Harley *et al.*, 1992; Mahadevan *et al.*, 1992)located on chromosome 19. The repeat expansion that a patient inherits will dictate to some extent the nature of the symptoms that will affect that person over their lifetime, while the progressive nature of the disorder means that the expansion of the repeat correlates with the presentation of certain disease symptoms. Generally the larger the repeat carried, the more severe the symptoms and the earlier the age of onset. The boundaries however, are not tightly defined, due possibly to a combination of patients presenting the disorder on diverse genetic backgrounds, being diagnosed at different stages of progression and in some cases retrospective recollection as to the order of symptom onset. Despite this, the degree to which the disease affects a patient has lead to the subdivision of the disorder into three groups (Table 1.2) with an indication of the range of the repeat sizes associated with each phenotype. The normal repeat size found at the DM1 locus is between 5 and 37 repeats, patients who are mildly affected have around 50 to 150 repeats, classical adult onset patients have between 100 and a 1000 repeats, while congenital patients have repeats into the thousands (Harper *et al.*, 1992; Mahadevan *et al.*, 1992).

Number of repeats	Phenotype	Clinical symptoms
50-150	Mild	cataracts
100-1000	Classical	myotonia, muscle wasting, testicular atrophy, premature balding, cardiac conduction defects
>1000	Congenital	mental retardation, facial displegia and jaw weakness, hypotonia

Table 1.2 The association between the DM1 phenotype and the size of the triplet repeat tract (Harper *et al.*, 1992; Mahadevan *et al.*, 1992).

It has already been mentioned that DM1 is distinct amongst muscular dystrophies in its multisystemic nature. Unlike other muscular dystrophies, muscles other than skeletal muscle are affected, as seen with the conduction defects in the heart. The smooth muscle is affected extensively resulting in a patient having problems swallowing, and within the digestive tract. The pattern of skeletal muscle degeneration is distinct too, with the distal muscles of the limbs predominately affected as well as many of the muscles controlling the face and jaw including those controlling the eyelids. As the disease progresses the diaphragm and intercostal muscles become affected, causing breathing difficulties, yet certain muscles types are not affected at all, such as the posterior calf and thigh muscles of the leg, suggesting a difference in muscles vulnerable to the mechanism of deterioration (Harper, 1998). In addition to these symptoms it is also well documented that DM1 patients exhibit an increased resistance to insulin (Hudson *et al.*, 1987; Krentz *et al.*, 1990; Moxley *et al.*, 1987).

1.3 Origins of the DM1 repeat

The incidence of DM1 is reported as 1 in 8000 in the Caucasian population (Harper, 1989), with a lower incidence of 1 in 18000 in the Japanese population. The disorder is rarely reported in African populations.

Anticipation through successive generations of DM1 families infers a correlation between increasing repeat size and decreasing reproductive fitness. This results in the termination of disease range pedigrees (Neville *et al.*, 1994). However, the continued incidence of DM1

within the population suggests the presence of a pool of pre-disease range alleles or protomutations.

The triplet repeat at the DM1 locus is polymorphic. Examination of the triplet repeat in DM1 alleles in the general population has lead to their allocation into pools or modes. Of the resulting three modes; 5 CTG repeats or (CTG)5 account for 35 - 40 %, (CTG)11, (CTG)12, and (CTG)13 account for a further 50 %. The final 10 % consists of a group with repeats in the range (CTG)19-30. There are no clearly defined sub groups within this third group, the increase in repeat number having already conferred some instability to this mode (Imbert *et al.*, 1993). This trimodal distribution was observed in African, European, Japanese and Native American populations (Brunner *et al.*, 1992). The (CTG)5 mode occurs most frequently in European populations (Brunner *et al.*, 1992; Davies *et al.*, 1992) followed by its incidence in the Asian population (Yamagata *et al.*, 1996). The second mode is most common in non Eurasian populations (Davies *et al.*, 1992; Watkins *et al.*, 1995; Zerylnick *et al.*, 1995).

The CTG repeat in the untranslated region (UTR) of *DMPK* is in total allelic association with an *Alu* deletion / insertion (-/+) polymorphism (Mahadevan *et al.*, 1993). A stretch of five *Alu* elements is associated with (CTG)5 and (CTG)19-30 alleles. This suggests that the repeats in the (CTG)19-30 group originate from a single expansion event from a (CTG)5 allele. That the (CTG)19-30 mode, the allelic pool from which the probands of DM1 pedigrees originate are from a genetic background in common with that of the stable (CTG)5, that is both modes in the first instance were subject to the same *cis* and *trans* acting factors, suggests that the instability present in the DM1 pedigrees is a consequence of the dosage of repeats and not the acquisition of another novel factor. The (CTG)11, (CTG)12, and (CTG)13 alleles are associated with two contiguous *Alu* elements (Neville *et al.*, 1994).

Deka *et al.* (1996), examined the DM1 CTG repeat tract and the associated *Alu* (-/+) polymorphism in samples collected from 16 ethnically and geographically diverse populations. They confirmed the (CTG)5 allele to be the most common allele in the majority of populations, but observed its absence from the Costa Rican and New Guinea highlander population. They also detected a (CTG)4 repeat allele, in a single American Samoan. This is the smallest reported CTG allele.

Although the (CTG)5 and (CTG)19-30 modes are almost always associated with the *Alu*+ element, Deka *et al.* (1996), report the presence of the (CTG)5 tract associated with an *Alu*- polymorphism in several populations including native Africans. They hypothesise that the

ancestral haplotype is (CTG)5 *Alu*+ and it gave rise to (CTG)5 *Alu*-. The larger (CTG) modes arose from either (CTG)5 *Alu*+ or (CTG)5 *Alu*-.

1.4 Somatic mosaicism, within and between tissues

The expanded allele of a DM1 patient is likely to appear on a genomic Southern blot as an expanded smear, though this depends on the age of the patient and the size of their repeat. In contrast to the normal range allele whose appearance is as a tight band (Jansen *et al.*, 1994; Joseph *et al.*, 1997; Lavedan *et al.*, 1993a; Tachi *et al.*, 1993; Tachi *et al.*, 1995b; Wöhrle *et al.*, 1995). However, the application of a more sensitive allele resolution technique, small pool (SP) PCR, (Martorell *et al.*, 1998; Wong *et al.*, 1995) has lead to the resolution of the smear into individual alleles, each of which has a different length of repeat. As the repeat gets larger, the more prone to expansion events it appears to become (Monckton *et al.*, 1995). Scrutiny of expanded DM1 repeats reveals expansion biased instability and variation of stability levels between different tissues taken from the same patient at the same timepoint (Jansen *et al.*, 1994; Joseph *et al.*, 1997; Lavedan *et al.*, 1993a; Tachi *et al.*, 1993; Tachi *et al.*, 1995b; Wöhrle *et al.*, 1995). It has been observed that the expanded repeat tract in skeletal muscle expands more quickly than in blood, but both tissues have a common lower boundary when subjected to small pool PCR, suggesting a common sized progenitor allele (Monckton *et al.*, 1995).

Examination of tissues; skeletal muscle, cardiac muscle, liver, lung, bowel, brain, skin and bone, of a 20 week old foetus with the CTG expansion at the DM1 locus (Lavedan *et al.*, 1993a), and thirty five tissue types obtained from congenital DM1 identical twins that had ~ (CTG)1500 repeats, who were born at 28 weeks and who both died 2 weeks after birth (Jansen *et al.*, 1994), showed small fluctuations of stability between tissues types. More importantly the same tissue type when compared between the twins revealed no difference in stability level at birth (Jansen *et al.*, 1994). To try to understand the repeat stability, or lack of stability observed between the same tissues from identical CDM twins it should be noted that identical or monozygotic twins separate at the earliest at the two cell stage, around 36 hours after fertilisation and at the latest just before the bilaminar disc develops into the primitive steak at around three weeks. It is most common however, for them to split at the early blastocyte stage around a week into development (Sadler, 1985). Comparison of blood samples between two further sets of monozygotic DM1 twins, one set with 180 and the other with 400 repeats showed no noticeable differences between the same tissues samples taken

from the twin pairs (Jansen *et al.*, 1994). The twin pair data suggests that some somatic differences may be laid down very early in embryonic development but that the repeat in developing twins will be subject to similar stability effecting factors as they develop. Precedence for this can be seen in the examination of day 10 embryos at the highly variable mouse tetranucleotide repeat *Hm-2* (Gibbs *et al.*, 1993). Gibbs *et al.* (1993) examined the result of 61 conceptions, studying the embryo and when possible, extraembryonic tissue. The extraembryonic tissue was usually the trophoblast. The mouse embryo or morulla splits from the body that will become the trophoblast when it reaches the 32 cell stage on the fourth day of development in both mice and humans. Examination of the *Hm-2* loci found mutations in 56% of the studied embryos, this rate was higher than that observed in live young, which was observed at 20%. Of the 46 trophoblast / embryo pairs studied five shared the same mutation. In these mice the mutation event must have therefore occurred within the first five cell divisions. So it is possible that a mutation event may occur at the effected disease loci before twins split. However, the tissues in each twin would be subject to the same rate of development and any specific factors, including environmental effects, acting on the different tissues would be the same. So, far from instability patterns being laid down in the one day to 21 day window defined by twin separation, the tissue specific differences could have occurred later in development, a consequence of common tissue specific factors. After all it is with the emergence of the primitive streak, laying foundations for brain, central nervous system and heart at around 3 weeks in humans and 7 days in mice, that the first evidence of tissue differentiation can be observed. So for tissue specific repeat mutations to have occurred prior to this, cells must be programmed early with their future function. Indeed, if the events contributing to tissue specific instability are laid down early in embryonic development, that is before differentiation, it could be that the rate of repeat expansion would be uniform across the tissue types.

Analysis of leukocyte DNA from disease range patients at timepoints ranging from neonates to those aged over 50 found that the diffuse band observed in genomic Southern blots of unstable disease range alleles wasn't present in the neonates. Indeed the band appeared to become more diffuse as the age of the patients analysed increased (Wong *et al.*, 1995) suggesting that somatic expansion events occur postnatally.

Initial comparison of the developmental timing of cell types (Sadler, 1985) located within tissues exhibiting phenotypes associated with DM1 found there to be no correlation with either the appearance or severity of the DM1 symptoms and the order of tissue development. Furthermore examination of the rate of cell proliferation and migration in adult rats and mice

using thymidine H^3 across static expanding and renewing tissue populations (Messier and Leblond, 1960), also shows no correlation with DM1 phenotype tissues or tissues known to show repeat instability.

1.5 Gametic instability

The sex of the transmitting parent and the size of the repeat expansion that they carry effects the size of the repeat inherited. Study of DM1 pedigrees found the transmitting grandparent most likely to be male (Bell, 1948) an observation confirmed in more recent studies (Lavedan *et al.*, 1993a). When a number of DM1 transmissions were studied the average size of repeat transmitted by fathers was found to be shorter than the average repeat length transmitted maternally (Lavedan *et al.*, 1993a). Fathers with a small repeat expansion tend to transmit large repeats to their offspring, while fathers with larger repeat sizes themselves do not tend to transmit as large jumps (Ashizawa *et al.*, 1994).

It should be noted that testicular atrophy is a common symptom in classic DM1 patients although it is also possible that there is a ceiling over which male gametes cannot expand. Comparison of blood and sperm samples of nine DM1 males showed the smeared expanded band characteristic of genomic Southern blot analysis present in the sperm samples. Sperm samples showed a higher degree of repeat mosaicism than the blood samples, the smear being indicative of the presence multiple repeat sizes (Jansen *et al.*, 1994). This observation is echoed in three of the DM1 patients studied by Monckton *et al.* (1995) and five out of the six DM1 patients from whom blood and sperm samples were examined using small pool PCR by Martorell *et al.* (2000). The expansion threshold for paternal DM1 transmission has been suggested at around 1000 repeats (Jansen *et al.*, 1994) although examples beyond this have been observed (Martorell *et al.*, 2000). Such expansion events however, appear to be rare and sperm with repeat expansions larger than 1000 CTG may be selected against as the sperm matures, or mechanisms may exist which act against large expansions in sperm. It may also be that the maturation time of sperm may not give it time to be subject to multiple small expansion events while the rare individuals who present sperm with repeats beyond the 1000 CTG repeat threshold may be subject to unusual expandability factors (Martorell *et al.*, 2000). Congenital myotonic dystrophy (CDM) was found to be transmitted almost exclusively maternally (Ashizawa *et al.*, 1994), although cases of paternal CDM transmission have been observed (Tachi *et al.*, 1997), contractions, when they occur, are more often transmitted by

12

fathers (Ashizawa *et al.*, 1994). Even when the CDM transmissions are excluded from the maternal mean, the average maternal transmission was still found to be larger.

These observations suggest that the phenotype a patient presents is related not only to inherited repeat size (Lavedan *et al.*, 1993a), but to other factors, such as the presence of neighbouring *cis* acting elements (Jansen *et al.*, 1994), or the effect of *trans* acting elements, it may be that certain individuals carry polymorphisms in genes involved in the mechanism of instability that make them more vulnerable to the effect of the triplet repeat (Wells, 1996).

1.6 *DMWD* and *SIX5*

As the investigation of the effect of the expanded CTG tract in the 3' UTR of *DMPK* in DM1 patients grew, the effect of the expansion on the genes closest to *DMPK* was also examined. *DMWD*, formerly referred to as *59* in humans (Shaw *et al.*, 1993) and *Dmr-n9* in mice (Jansen *et al.*, 1995; Jansen *et al.*, 1992), is located ~1.1 kbp upstream of the start of transcription of *DMPK* (Figure 1.2) in both humans and mice. This gene contains five exons and encodes a 650 amino acid protein (Jansen *et al.*, 1992) which includes two inner domains that show significant homology to a WD motif (Jansen *et al.*, 1995). The WD amino acid sequence is a conserved motif found in a family of proteins involved in signal transduction or cell regulation. Jansen *et al.* (1992) observed the presence of two weak adenylation sites adjacent to the genes and hypothesised contiguous transcription with the adjacent *DMPK* gene, arguing that factors affecting the *DMPK* transcription could have an affect on flanking DNA. The observation that normal Dmwd expression is most prominent in the brain and testis (Jansen *et al.*, 1995) and that DMWD is expressed in the heart, brain, liver, kidney spleen and testis in humans has lead to the suggestion that the expansion of the DM1 repeat is polygenic in its effect. *SIX5*, formerly known as *DMAHP, (MYOTONIC DYSTROPHY LOCUS ASSOCIATED HOMEOBOX PROTEIN)* ~ 1.2 kbp downstream from the final *DMPK* exon (Boucher *et al.*, 1995; Heath, 1997) (Figure 1.2) in both humans and mice. A homeobox gene encodes a conserved domain of ~60 amino acids, a homeodomain. This "homeodomain" can bind to a specific DNA sequence (Engelkamp and van Heyningen, 1996). *SIX5* has been identified as a member of the growing *sine oculis related homeobox (Six)* subfamily named after the first member identified, the *Drosophila* eye development gene *sine oculis (so)*. Indeed *Six5* (-/+) and (-/-) mice have been generated by two groups and both models exhibit ocular cataracts, a disease phenotype manifested in late onset patients (Klesert *et al.*, 2000; Sarkar *et al.*, 2000a). The *Six5* -/-, mice reported by Sarker *et al.* (2000a), have been reported as being sterile
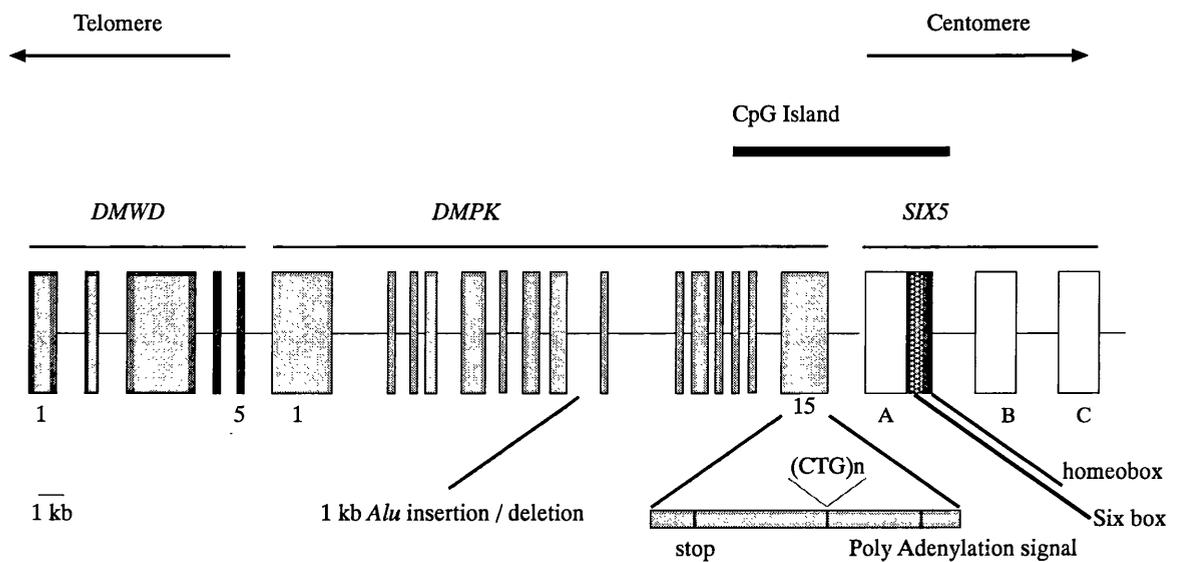
Figure 1.2 Genomic structure of DM1 region

Schematic diagram of *DMPK* and the flanking genes *DMWD* and *SIX5*. Indicated is the location of each exon, the expanded triplet repeat, the *Alu* insertion/ deletion polymorphism and the *SIX5* homeobox and six box.

demonstrating a role for Six5 in gametogenesis or embryogenesis (Sarkar *et al.*, 2000b). They also report muscle wasting and suggest that decreased expression of Six5 may contribute to DM1 skeletal myopathy.

1.7 The effects of the triplet repeat expansion on flanking DNA

As the *DMPK* trinucleotide repeat is located outside the open reading frame of the protein coding region it is unlikely that it influences the actual protein structure in the manner of the polyglutamine tracts of CAG triplet repeat diseases (Cummings and Zoghbi, 2000). As yet no patient with DM1 symptoms has been reported with a point mutation within *DMPK* that would knockout its function. What the resulting phenotype of such a mutation might be is therefore the subject of speculation. To investigate the effect of the loss of the DMPK protein, *Dmpk -/-* mouse models have been generated (Figure 1.3). Such knockouts (Jansen *et al.*, 1996; Reddy *et al.*, 1996), were viable both as heterozygotes and unexpectedly as homozygotes but did not result in the manifestation of all of the disease associated symptoms. The Jansen *et al.* (1996) (-/-) mice exhibited minor muscle fibre changes in the head and neck muscle fibres, but no myotonia. Although the effects of the *Dmpk* (-/-) mutation appeared to be minor when looking at the gross mouse, observation of Ca^{2+} homeostasis at the cellular level in the Jansen *et al.* (1996) (-/-) mice found voltage-dependent L-type Ca^{2+} and Na^{2+} channels to exhibit smaller and slower Ca^{2+} responses to acetylcholine or high external K^+ stimuli (Benders *et al.*, 1997). While more recently abnormality in the sodium gating channels in the myocytes of the Reddy *et al.* (1996), *Dmpk -/-* mice have been observed to resemble the sodium channel abnormalities presented by DM1 patients (Mounsey *et al.*, 2000).

15

A   Human DM region

Chr 19

(CTG)n

DMWD   DMPK   SIX 5

CpG island

DM55 Gourdon et al. (1997)

(CTG)55

DMWD   DMPK   SIX 5

45 kb

Dmt-162 Monckton et al. (1997)

(CTG)162

~0.8 kb

B

Murine DM region

Chr 7

CTGCTGCAGCAGCTG

Dmwd   Dmpk   Six 5

CpG island

Dmpk knockout Jansen et al. (1996)

CTGCTGCAGCAGCTG

Dmwd   Hygro B   Dmpk   Six 5

CpG island

Dmpk knockout Reddy et al. (1996)

CTGCTGCAGCAGCTG

Dmwd   Neo   Dmpk   Six 5

CpG island

Figure 1.3 DM1 related transgenes

Schematic diagram of A. the human DM1 region and derived transgenes and B. the homologous murine region and the related murine models.

16

The Reddy *et al.* (1996*)*, mice presented a progressive skeletal myopathy increased muscle fibre degeneration, fibrosis, a 50% loss of force generation and ultrastructural changes. They share some pathological features with DM1, suggesting that *Dmpk* is functionally necessary to maintain the structural integrity of skeletal muscle fibres and normal force generation (Reddy *et al.*, 1996). Berul *et al.* (1999), reported that the Reddy *et al.* (1996), *Dmpk* -/- mice developed cardiac conduction defects while the *Dmpk* +/- mice defects developed a similar condition. This cardiac defect resembled the heart defects observed in DM1 patients (Berul *et al.*, 1999).

Over expression of a *DMPK* transgene in mice revealed an expression pattern in common with that of *Dmpk* but 20 transgenic copies were required to increase *DMPK* expression to 5 – 10 times the levels of *Dmpk* expression (Jansen *et al.*, 1996). The mice appeared fundamentally normal, although histopathological analysis of the heart revealed a striking dysmorphology which included patches of degeneration, yet cardiac function appeared normal, so any correlation this effect may have with the cardiac conduction defects presented by DM1 patients was not apparent. It should be noted before drawing conclusions from the lack of manifestation of disease phenotype that neither of the mouse models for *DMPK* repeat expansion (Figure 1.3), (Gourdon *et al.*, 1997; Monckton *et al.*, 1997) exhibit disease phenotypes although variation in the repeat is observed both somatically and gametically.

It is known that expanded CTG triplet repeat blocks from DM1 form the strongest known natural nucleosome positioning elements (Wang *et al.*, 1994). DNA fragments containing 75 and 130 repeats respectively were observed to be 6 and 9 times stronger in nucleosome formation than that of the somatic 5S RNA gene of *Xenopus borealis*, previously one of the strongest known positioning elements. The consequence of the expanded repeat tract is a region of condensed chromatin and also the loss of a DNAse I hypersensitivity site, located 3' of the triplet repeat (Otten and Tapscott, 1995), suggesting that access to transcription factor sites might also be effected by structural changes to the region. Examination of the methylation pattern of the CpG island 3' to *DMPK* revealed no sex related differences in the methylation pattern (Steinbach *et al.*, 1998). However, a difference in methylation status has been observed in different classes of DM1 patients. The CpG island in the classic cases of DM1 examined were reported to be normal while the same region in all of the CDM cases studied were found to be hypermethylated.

1.8 RNA effects and pathology

The role of *DMPK* and the mechanisms by which the expansion of the triplet repeat tract leads to the DM1 phenotype are not known. Speculation as to what the pathogenic mechanism might be has lead to the proposal that the expansion leads to a failure in the processing and / or export of the mRNA transcripts of *DMPK* and that the *DMPK* expanded repeat tract, (mutant), interacts with the transcript of the *DMPK* allele that does not contain the expansion, (normal), (Fu *et al.*, 1993; Hamshere *et al.*, 1997; Krahe *et al.*, 1995; Roses *et al.*, 1996; Sabourin *et al.*, 1993; Wang *et al.*, 1995) and / or the transcripts of the flanking genes *DMWD* (Alwazzan *et al.*, 1999; Hamshere *et al.*, 1997) and *SIX5* (Hamshere *et al.*, 1997; Klesert *et al.*, 1997; Thornton *et al.*, 1997).

In the first instance, investigation of the levels of transcription of the mutant and the normal alleles of *DMPK* produced conflicting results. For example Wang *et al.* (1995) observed that the level of *DMPK* polyA RNA was reduced for both the mutant and the normal transcript when compared to a non disease range control. In contrast Fu *et al.* (1993), observed a lack of coequal expression of the mutant and normal *DMPK* transcript levels. The transcription level of the normal allele was uneffected, however, a reduction of mutant levels was detected and this reduction increased as the size of the repeat increased. Krahe *et al.* (1995), detected equal levels of each *DMPK* allele in the analysed pool of heteronucleic RNA. Analysis of the *DMPK* levels in the pool of mRNA revealed a reduction in the level of mutant *DMPK* detected. Again this reduction increased with the increasing repeat size. A similar observation was made by Davis *et al.* (1997) and Hamshere *et al.* (1997). Sabourin *et al.* (1993), conversely reported elevated levels of *DMPK* mRNA transcript in the tissues and cell lines that they examined and attributed this to an elevated level of mutant *DMPK* mRNA.

The data presented in these studies consists of analysis of transcript levels of the mutant and normal allele in the messenger RNA fraction and / or the heteronucleic RNA fraction. The RNA was extracted from either the whole cell (Krahe *et al.*, 1995; Wang *et al.*, 1995), or separated into nuclear and cytoplasmic fractions (Alwazzan *et al.*, 1999; Hamshere *et al.*, 1997). The RNA has been derived from a number of sources, including congenital patients (Sabourin *et al.*, 1993) and classic onset DM1 patients (Fu *et al.*, 1993; Krahe *et al.*, 1995), as well as DM1 patient derived cell lines (Alwazzan *et al.*, 1999; Davis *et al.*, 1997; Hamshere *et al.*, 1997; Krahe *et al.*, 1995). It is now clear that congenital cases have different symptoms to classic onset patients.

The methods used to analyse allele levels were also diverse and will have varied in their sensitivity. They included northern blotting (Davis *et al.*, 1997), RNA fluorescent *in situ*

18

hybridisation (Taneja *et al.*, 1995), quantitative multiple fluorescent (QMF) PCR (Wang *et al.*, 1995) and RT-PCR (Fu *et al.*, 1993; Hamshere *et al.*, 1997; Sabourin *et al.*, 1993) including allele specific RT-PCR, (Alwazzan *et al.*, 1999; Klesert *et al.*, 1997; Krahe *et al.*, 1995). This vast variation in the methods used to analyse the transcript levels goes some way to explain the conflicting results.

Current thinking about the expression of and the export of the mutant and normal *DMPK* allele is based on RT-PCR carried out on mRNA prepared from separated nuclear and cytoplasmic fractions (Hamshere *et al.*, 1997) and allele specific expression of *DMPK* (Krahe *et al.*, 1995). These studies have shown the level of transcription is not effected by the expanded allele but that processing and maturation of the transcript is (Krahe *et al.*, 1995). The rate that *DMPK* transcripts accumulate for both normal and myopathic controls when compared to accumulation of levels of both the normal and expanded allele of DM1 saw a small decrease in *DMPK* RNA levels amongst the pool of total RNA and as the myopathic control showed a similar decrease, this fluctuation did not appear to be disease specific (Taneja *et al.*, 1995). All of the RNA under examination was extracted from biopsied muscle. However, dramatic decreases of both the mutant and normal *DMPK* mRNA transcript was observed and this was disease specific. This suggested that the expanded repeat contributes to the nuclear retention of both transcripts. Meanwhile RNA *in situ* hybridisation revealed that the expanded *DMPK* transcript is retained within the nucleus and forms distinctive RNA foci (Taneja *et al.*, 1995). It might be assumed that the reduction of *DMPK* mRNA in DM1 patients would result in a reduction of the DMPK protein product and that this loss contributes to the symptoms manifested by a DM1 patient.

1.8.1 The effect of the *DMPK* expanded triplet repeat on flanking genes

Having explored the consequences of the expansion on the transcription and translation of *DMPK*, the effect that the expanded repeat may have on the outlying genes, *DMWD* which is located upstream of *DMPK* and *SIX5* which is located downstream of *DMPK* (Figure 1.2) was studied.

Hamshere *et al.* (1997), analysed *SIX5* levels in RNA fractions prepared from either the nuclear or cytoplasmic fraction of a DM1 derived cell line. They observed there to be no difference in the level of *SIX5* RNA present in the two fractions. Klesert *et al.* (1997) and Thornton *et al.* (1997), also analysed the transcription level of *SIX5*. They examined levels in both the normal and mutant alleles using allele specific RT-PCR. This revealed a reduction in

the level of mutant transcript that was 2-4 times less than that observed for the normal allele (Klesert *et al.*, 1997). While Thornton *et al.* (1997) observed that as the repeat expansion increased the *SIX5* transcript level dropped. The DNase I hypersensitive site that was reported by Otten and Tapscott, (1995), was found to contain an enhancer element that regulated the transcription of *SIX5* and the disruption of this site by the expanded triplet repeat reduced the level of transcription (Klesert *et al.*, 1997; Thornton *et al.*, 1997). Klesert *et al.* (1997) and Thornton *et al.* (1997) have demonstrated that the reduction of *SIX5* transcript levels is because of a reduction in transcription initiation and not as in *DMPK* a problem with transcript processing. This explains the co-equal levels of *SIX5* transcript observed in the nuclear and cytoplasmic fractions observed by Hamshere *et al.* (1997). Allele specific RT-PCR analysis of the same DM1 cell lines reported by Hamshere *et al.* (1997) also revealed a reduction in the level that the *SIX5* was transcribed from the mutant allele (Alwazzan *et al.*, 1999), and that the larger the repeat expansion the greater the reduction in mutant transcript level, that is the cell lines containing the larger expansions, 3 – 5.5 kb had a greater effect reducing expression levels than the repeats between 240 bp and 1.2 kb in size.

To assess the effect of the CTG expansion on the level of transcription of *DMWD,* Alwazzan *et al.* (1999), designed allele specific primers for both the mutant and normal allele of *DMWD.* Thus allowing the level of each of these transcripts to be monitored both in the nucleus and the cytoplasm. A series of DM1 patient derived fibroblast cell lines containing a selection of repeat lengths were used for the study. Comparison of the level of normal transcripts versus mutant transcripts isolated from the nucleus found a reduction in the level of mutant transcript to between 20 – 50% of that of the level of the normal transcript. However, no correlation between the reduction of *DMWD* transcription levels and repeat length was observed. Comparison between *DMWD* transcript levels in cytoplasmic and nucleic fractions found there to be a reduction in the levels of the mutant allele. This suggests that the level of *DMWD* transcription is influenced at the initiation stage and this is in some way effected by the size of the triplet repeat expansion, influencing the cytoplasmic levels of the transcript.

1.8.2 Consequences of the expanded repeat in *DMPK* transcripts
It has been observed that the expansion of the mutant *DMPK* CTG repeat results in a reduction of transcript levels of *DMPK* detectable in the cytoplasm (Hamshere *et al.*, 1997). So what could the mechanism of transcript retention in the nucleus involve?

Mankodi *et al.* (2000) generated several mouse lines using a transgene containing a genomic fragment of the human skeletal actin gene. They introduced a tract of either (CTG)5 or (CTG)250 repeats between the stop codon and the polyadenylation site of the transgene. Thus mimicking the location of the expanded repeat in DM1.

One of the long repeat carrying transgenic lines, LR32a was analysed because the transgenic repeat was expressed at a high level and silenced infrequently. These mice gained weight normally and histological analysis of non-muscular tissue showed this to be normal also. However, the mortality rate at 44 weeks had increased from 5% in non transgenic mice or transgenic lines generated with the (CTG)5, repeat to 41%. The LR32a mice showed abnormal hind limb posture after a period of inactivity and presented myotonia as young as 4 weeks although the muscles appeared histologically normal. Yet mice expressing the (CTG)250 repeats did develop histologically defined myopathy not observed in (CTG)5 mice. The long CUG transcripts were observed to be retained in the nucleus in discrete foci reminiscent of those observed by Taneja *et al.* (1995) in the myoblasts of DM1 patients. The presence of the expanded CUG repeat would appear to be enough to result in this retention. This in itself begs questions about what the mechanism of retention may be. What is the nature of the structure that is retained in the nucleus? It has been observed *in vitro* that CUG tracts can form hairpin structures as they are transcribed (Tian *et al.*, 2000). Examination of the structures formed by a selection of CUG repeats *in vitro*. Study of transcripts of (CUG)35, (CUG)69, and (CUG)140 found the preferred structure formed to be a simple hairpin containing an extended stem (Figure 1.4) rather than the alternative looping structures. The retention of the transcript could be accounted for by the assumption of such a structure by the expanded CUG tract. Koch and Leffert (1998), suggested that the giant hairpin structures might sterically block RNA export through nuclear pores. Successfully exported *DMPK* hnRNA's are 9-25 nm, when measured cross sectionally through the diameter. However, as the repeat expands the steric demands of the hairpin may result in a structure that is too large to be exported out of the nucleus. Although the presence of hairpin structure *in vivo* has not yet been confirmed, the consequences of such a structure might explain the multiple disease phenotypes of DM1.

A

```
        U        U        U        U        U        U        U
5'  C U G C   G C    G C    G C    G C    G C    G C
              | |    | |    | |    | |    | |    | |       G
3'  G U C G   C G    C G    C G    G C    G C           C
        U        U        U        U        U        U        U
```

B

```
          U        U        U        U        U        U        U
      C   G C    G C    G C    G C    G C    G C
      G   | |    | |    | |    | |    | |    | |       G
          C G    C G    C G    C G    G C    G C       C
      U        U        U    U    U  U        U        U
                             C    G
                             G    3'
                             U
                             C
                             5'
```
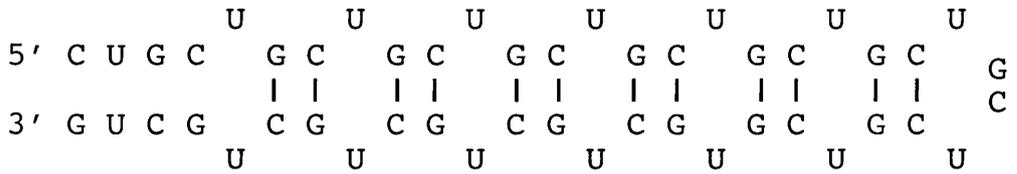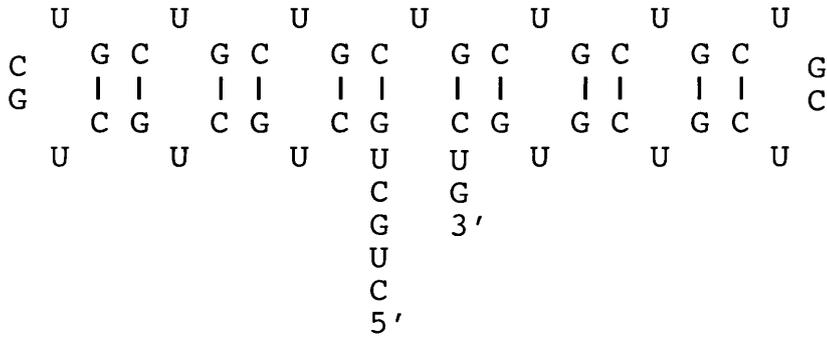
Figure 1.4 Secondary structures adopted by expanded CUG repeats.
Examples of secondary structures that could be adopted by CUG repeat tracts.
A. Simple hairpin with its extended stem
B. An alternative multiloop structure.

22

It has been hypothesised that the hairpin structures act as protein sinks removing some cellular proteins from their normal functions which may include RNA processing or export (Timchenko *et al.*, 1999). The removal of a protein or proteins with multiple functions could in some way account for some of the symptoms of the DM1 disease phenotype. As may the retention of the CUG transcripts within the nucleus. The accumulation of an "overdose" of transcript may in some way interfere with the balance within the nucleus and disrupt the normal function of proteins located there (Tian *et al.*, 2000).

The CUG binding protein (*CUG-BP*) or *hNab50* is observed to bind preferentially to CUG repeats, although not necessarily expanded ones (Timchenko *et al.*, 1999). It is known that *CUG-BP* accumulates in the nucleus and this could be explained by the presence of the repeat expansion. *CUG-BP* had been found to bind to the pre mRNA transcript of human cardiac troponin T (*cTNT*) where it regulates alternative splicing. It occurs in multiple isoforms and its function can be reliant on its phosphorylation status. The theory that the increase in potential binding sites resulting from the increasing number of repeats results in a "mopping up" effect and consequential loss of *CUG-BP* to its normal functions is confounded by the observation that *CUG-BP* binds to single stranded repeats (Michalowski *et al.*, 1999). Michalowski *et al.* (1999), used electron microscopy to observe repeat tracts generated *in vitro*. These CUG hairpins are double stranded structures uniformly paired along the repeat tract, without kinks, bubbles and bends restricting the potential binding region to the base of the hairpin. *In vivo* however, *CUG-BP* could potentially bind to the repeat tracts as they are generated.

The disruption of this protein or one like it with multiple functions may account for the variety observed in the disease phenotypes manifested in DM1 patients, (Phillips *et al.*, 1998). Alternatively if the *in vitro* observations of Michalowski *et al.* (1999), also occur *in vivo* and the repeat binds into a double stranded structure unaffected by *CUG-BP*, then this protein is not responsible, or solely responsible for the disease phenotype and it may act in tandem with a protein which recognises double stranded tracts. Tian *et al.* (2000), report that CUG repeats bind to the double stranded RNA binding domain of PKR, the double stranded RNA activated protein kinase (Tian *et al.*, 2000). Their *in vivo* studies have revealed that the CUG repeat tract must be in excess of (CUG)15 for binding to occur and that the affinity of PKR to the repeat tracts increases as the length of the repeat tract increases. It is possible that the retention or activation of PKR may have subsequent consequences that may result in DM1 disease symptoms. Whilst Miller *et al.* (2000), were investigating double stranded RNA binding proteins they isolated proteins homologous to the *Drosophila* muscleblind proteins (Miller *et*

23

al., 2000). These proteins are required for the terminal differentiation of muscle and photoreceptor cells. Both the eye and an extensive number of muscles are affected by DM1, suggesting that such proteins may account for some of the DM1 phenotype. The expression of muscleblind proteins is activated during mammalian myoblast differentiation and they have been shown to aggregate as nuclear foci in DM1 cell lines.

There has also been one case reporting an expanded CTG repeat causing *trans* RNA interference (Sasagawa *et al.*, 1999), where the interaction of an expanded (CTG)140 transcript with the (CAG)35 repeat of the TFIID binding protein results in an abnormal double stranded structure. This could result in not only nuclear retention of the transcript but affect post-transcriptional processes and result in a transcript unviable for nuclear export.

1.9 Transgenic mouse models of triplet repeat disease

There are many problems when it comes to studying the progression of the triplet repeat diseases. Access to a selection of tissue types from an individual is limited, indeed the means to examine the progression of the disorder in a range of tissues is restricted to post mortem tissue (Jansen *et al.*, 1994; Joseph *et al.*, 1997; Wohrle *et al.*, 1995). The observation of the repeat instability over time in DM1 is limited to blood samples (Lavedan *et al.*, 1993a; Martorell *et al.*, 2000; Massari *et al.*, 1995; Tachi *et al.*, 1993; Tachi *et al.*, 1997), biopsied muscle (Massari *et al.*, 1995; Tachi *et al.*, 1993), which is limited due to the invasive nature of the procedure and semen samples (Martorell *et al.*, 2000; Massari *et al.*, 1995), the latter is a source which becomes limited as the disease progresses. Other factors which confound investigation include the difficulty in determining the initial size of the disease range repeat. This is especially difficult as patients are diagnosed at a range of ages and at different stages of disease progression. They have been subject to a variety of environmental factors and knowledge of disease progression is limited as even within families there will be variation in the genetic background and thus variation of the *cis* and *trans* acting factors that are acting upon the repeat.

 The study of intergenerational repeat variation is confounded by the low numbers of offspring in affected families. Large timepoints between generations and limited diagnosis of earlier less affected generations make it difficult to follow repeat transmission patterns. That parental age has been observed to effect only the size of the transmitted Huntington repeat (Duyao *et al.*, 1993) and DRPLA repeat (Sato *et al.*, 1999) may reflect limited pedigree data.

24

The lack of consistent sources of triplet repeat tissues has lead to the development of both murine and cell models. There are arguments against the use of mouse models after all, how appropriate is it to use the mouse, with its short lifespan to study a late onset disorder? Factors affecting repeat stability and thresholds above which a triplet repeat tract might expand may vary between species or indeed between strains of model organisms. However, the short lifespan lends itself to the study of successive generations as well as the study of mature individuals. There is also ease of access to multiple tissues of known progenitor repeat size at any time point. Use of inbred strains means that the expanded repeat is presented on a uniform genetic background and variation of the repeat length can be the only difference between individuals as even environmental variation is within the control of the experimenter. A number of murine models have been developed to investigate the effect of either an expanded CAG·CTG triplet repeat tract or any of a number of triplet repeat disorders in the hope of increasing understanding of somatic and / or intergenerational stability, the mechanisms involved and any variation in the mechanisms of stability that may exist between the two. The transgenic models split into two groups, those which involve transgenes based on the disease associated expanded repeat tract and differing amounts of flanking sequence which have integrated randomly into the genome (Bingham *et al.*, 1995; Burright *et al.*, 1995; Goldberg *et al.*, 1996; Gourdon *et al.*, 1997; Ikeda *et al.*, 1996; La Spada *et al.*, 1998; Mangiarini *et al.*, 1997; Monckton *et al.*, 1997; Sato *et al.*, 1999) and more recently those generated by the recombination of the expanded repeat tract into the murine homologue of the disease affected gene (Lorenzetti *et al.*, 2000; Shelbourne *et al.*, 1999; Wheeler *et al.*, 1999).

Not all of the generated mouse models show detectable instability. These include two transgenic lines generated to study SBMA containing either 45 or 24 CAG repeats derived from the human androgen receptor (AR) (Bingham *et al.*, 1995), cDNA derived Huntington constructs of 82 CAG repeats (Goldberg *et al.*, 1996) and an expanded CAG sequence (Ikeda *et al.*, 1996). These models did not exhibit intergenerational or somatic instability or any evidence of disease phenotypes. All of these stable models contain what could be considered short repeat tracts although many of these repeat tracts may manifest symptoms in affected patients. It should be noted that these models contain multiple transgenic integrants, this factor and the sensitivity of methods used to detect the instability may make subtle changes in repeat size difficult to detect.

It might be that short repeat tract models do not reach a possible instability threshold during the lifetime of the mouse and that the level of instability observed is a function of the "dosage of repeats". Indeed this point is illustrated by the three "knock in" models of CAG expansion

25

at the murine HD loci, which contain an expanded repeat of (CAG)48, (CAG)90 and (CAG)109 generated by Wheeler *et al.* (1999). Minimal, (less than 2%), intergenerational instability was observed in the (CAG)48 line, while the frequency of mutations associated with the expanded repeat in the (CAG)90 and (CAG)109 is over 70% (Wheeler *et al.*, 1999). However, among the murine models of triplet repeat expansion which do exhibit gametic instability (Burright *et al.*, 1995; 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; Wheeler *et al.*, 1999) and somatic instability (Gourdon *et al.*, 1997; Mangiarini *et al.*, 1997; Monckton *et al.*, 1997; Sato *et al.*, 1999; Shelbourne *et al.*, 1999; Wheeler *et al.*, 1999) is the DM1 transgenic model, DM 55 (Gourdon *et al.*, 1997).

The DM 55 transgene consists of an expanded CTG repeat, derived from a DM1 patient, contained within ~ 45 kb of sequence that includes the coding sequence for *DMPK*, *DMWD* and *SIX5*. Despite the relatively short repeat tract (CTG)55, moderate intergenerational and somatic instability is observed, suggesting that the somatic instability observed is a function of the flanking DNA and when the transgene contains a minimal amount of flanking sequence (Ikeda *et al.*, 1996; Monckton *et al.*, 1997) a consequence of the integration site or other stability affecting factors. The consequence of such factors can be further observed in the somatic instability profiles of the R6 transgenes of Mangiarini *et al.* (1997), as instability is observed in the mature mice of the R6/1, R6/2 and R6/5 lines but not the R6/0 line (Bates *et al.*, 1997). Indeed between the unstable lines the instability profile, which was visualised using Genescan® software, varied. The SCA type 1 CAG repeat in normal range individuals contains a CAT repeat which is not present in expanded alleles and is thought to confer a stabilising effect (Jodice *et al.*, 1994). A point mutation in a transgenic repeat tract may confer the same effect. More likely to influence stability of the transgene is the position of the integration site and the DNA packaging of the transgene which may result in it being silenced yet still allowing it to be inherited (Clark, 1997). While Brock *et al.* (1999) observed that the GC content flanking the 10 triplet repeat disease loci examined correlated with the degree of repeat expandability observed there. That it is the variation in the integration sites which effects the observed instability in the same transgenic repeat tract can be seen when investigating different lines generated with the same transgene (Bates *et al.*, 1997; Monckton *et al.*, 1997).

Examination of the transmission patterns of triplet repeat murine models found the repeat size difference of parent offspring transmissions to be smaller than those observed in affected families (Burright *et al.*, 1995; 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; Wheeler *et al.*, 1999). A dosage of repeat effect could also be observed in the transmission of one of the *Dmt* transgenes. Line *Dmt*-A, which contains multiple copies of the *Dmt*-A transgene, was observed to transmit deletions of up to 38 repeats (Monckton *et al.*, 1997). The effect of the age of the transmitting parent on the size of the repeat transmitted has been studied in several models of triplet repeat instability and again sex related differences have been observed, Lorenzetti *et al.* (2000) observed that paternal transmissions of their targeted *Sca* type 1 model showed a low level of instability on which the age of the transmitting parent had little effect, while the deletions associated with the majority of the maternal transmissions increased with the age of the mother (Lorenzetti *et al.*, 2000), while Kaytor *et al.* (1997) also observed that as the age of the repeat transmitting mother increased so did the size of the deletion transmitted. They also examined the repeat size in oocytes isolated from mice of increasing age and found the same correlation, leading them to conclude that the transgenic repeat instability event occurs after meiotic recombination and prior to oocyte fertilisation.

1.10 Mismatch repair genes

Study of *E. coli* strains found those lacking one of 4 mutator genes, *mutS, mutL, mutH,* or *mutU*, resulted in an elevated level of spontaneous mutations (Modrich and Lahue, 1996). These mutator genes were thought to play a role in DNA mismatch repair, (MMR) a mechanism by which the integrity of the genome is maintained. A summary of the known roles of these mutator genes is shown in table 1.3 (Buermeyer *et al.*, 1999). Homologues for two of the mutator genes, *mutS* and *mutL* have been found in eukaryotes, both in yeast and mammals, although not all of these are thought to have a role in mutation avoidance. The homologues, which consist of heterodimers, are summarised in table 1.4 (Buermeyer *et al.*, 1999).

Eukaryotic homologues of *mutH* and *mutU* have not been identified. However, there is evidence to suggest that proliferating cell nuclear antigen (*PCNA*) has a role in MMR. It is known to interact with either human or yeast *MutS* and *MutL* homologues (Johnson *et al.*, 1996; Kokoska *et al.*, 1999). Mutations in *PCNA* have been observed to disrupt MMR, resulting in simple frameshifts and its role in MMR is prior to or at the point of excision (Umar *et al.*, 1994b), and a role has been hypothesised for *PCNA* which involves differentiating between the template and the nascent strand (Buermeyer *et al.*, 1999; Johnson *et al.*, 1996; Kokoska *et al.*, 1999) Of the mutator gene homologues, defects in at least five, *hMSH2, hMSH6, hMLH1, hPMS1* and *hPMS2* (Akiyama *et al.*, 1997), (Liu, 1996; Peltomaki and Vasen, 1997) have been implicated in the incidence of HNPCC (Lynch and Smyrk, 1996). A person with a mutation in one of these genes also has an increased risk of suffering from Muir-Torre syndrome and Turcots syndrome (Paraf *et al.*, 1997), which also involve colorectol cancers, as well as developing tumours in the endometrium, ovary, stomach and small intestine, (Lynch *et al.*, 1997; Peltomaki and de la Chappelle, 1997). However, of the MMR mutations observed in mismatch repair kindreds, the majority, are found in either *MSH2* or *MLH1*, which are thought to play a primary role while mutations in *PMS1*, *PMS2* and *MSH6* are apparently rarer (Liu, 1996).

Mutator gene	Role
mutS	An ATPase- which acting as a homodimer binds to both base / base mismatches and small insertion / deletion loops, that have not been connected by the replicating polymerase.
mutL	Couples mismatch recognition and downstream MMR events: • *mutL* homodimer forms complex with *mutS* complex- enhances ATP hydrolysis dependant translocation. • *mutL* stimulates *mutH* endonuclease activity in an ATP dependant manner • *mutL* required to load *mutU* helicase onto the site of a *mutH* induced nick
mutH	An endonuclease: introduces a nick in the nascent strand at hemi methylated GATC sequences allowing discrimination between the template and the newly replicated strand.
mutU	Helicase II/ UvrD- facilitates DNA unwinding.

Table 1.3 Summary of the role of the MMR mutator genes of *E.coli*

E.coli	*H. sapiens*	heterodimer
mutS	MutSα MutSβ	(hMSH2/hMSH6) (hMSH2/hMSH3)
mutL	MutLα MutLβ	(hMLH1/hPMS2) (hMLH1/hPMS1)
mutH	Not known	
mutU	Not known	

Table 1.4 Summary of the known human homologues of the MMR mutator genes of *E.coli*.

1.10.1 Mismatch repair genes and microsatellite instability

Mismatch repair genes 'proof read' the DNA, correcting errors in base pairing, unpaired regions and removing unusual DNA structures, such as hairpins (Umar *et al.*, 1994a; Yu *et al.*, 1995). However, it is now emerging that mismatch repair gene homologues also have other cellular functions including roles in transcription coupled repair (Leadon and Avrutskaya, 1997; Mellon and Champe, 1996), recombination (Chen and Jinks-Robertson, 1999; Rayssiguier *et al.*, 1989; Selva *et al.*, 1995) and possibly cell cycle regulation (Narayanan, 1997).

The absence or loss of function of these MMR genes result in mismatch errors accumulating over time with repair deficient yeast cells having been observed to accumulate mutations at a rate 700 times that of a non mutated yeast (Lynch and Smyrk, 1996).

Accumulation of such errors can eventually lead to the cancerous mutations observed in HNPCC patients. The repair of DNA errors acquired during the replication, repair and recombination of DNA is a balance between the survival of the organism and the acquisition of advantageous mutations. There are regions within the genome, which appear more vulnerable to mutation than others. These regions or "hot spots" are often associated with repeated DNA sequences, for example, inverted, mono, di, tri, and tetra nucleotide repeats and also some repeat tracts of higher magnitude, yet the trinucleotide repeats appear even more unstable than expected (Sinden, 1999). The migration of linearised triplet repeat sequences through acrylamide gels at a rate slower than that expected suggests the adoption of alternative structures by the triplet repeat tracts. It has been observed that several different hairpin structures can form when the triplet repeat tract is unpaired (Pearson *et al.*, 1998) and that DNA polymerase has difficulty passing through triplet repeat tracts, and that this itself is dependent on the repeat motif (Kang *et al.*, 1995; Ohshima *et al.*, 1996; Parniewski *et al.*, 2000). This leads to suggestions that the stalling of the DNA polymerase could lead to primer template misalignment resulting in the formation of alternative DNA structures, whose "repair" from the genome could result in the expansion and deletion events observed at expanded triplet repeat tracts.

Observation of a selection of CAG·CTG triplet repeat tract containing plasmids that were transformed into *Escherichia coli* revealed that all of the observed lengths of repeat tract were

subject to small deletions, that is deletions of less than 8 repeats when the mismatch repair pathway was inactivated while an active MMR pathway resulted deletions larger than 8 repeats. It was also observed that the larger the CAG·CTG triplet repeat tract was the larger the deletions observed (Jaworski *et al.*, 1995). To further confirm this observation CAG·CTG triplet repeat tract containing plasmids which varied in length and purity were transformed into *E.coli* strains that were isogenic except when a mismatch repair protein coding gene was mutated (Parniewski *et al.*, 2000), thus controlling the variation of *cis* and *trans* acting factors. It was observed that the loss of MMR proteins reduced the level of instability observed in the repeat tracts. The degree to which the CAG·CTG triplet repeat tract was vulnerable to deletions increased with the length of the tract as it did the less pure the tract under examination was. Deletions greater than 8 repeats were almost always observed when the repeat tracts were greater than 100 only. While the repeat tracts of less than 100 in the Parniewski *et al.* (2000), model, in common with other models both in *E. coli* (Schumacher *et al.*, 1997) and *Saccharomyces cerevisiae* (Schweitzer and Livingston, 1997) exhibited low levels of instability, +/- 1 repeat when MMR was functional. The observation that MMR promotes large deletions over a certain repeat tract threshold suggests at least 2 pathways promoting long or short deletions (Schumacher *et al.*, 1997) and the larger the repeat tract the greater the chance of the formation of alternative structures which may be edited out as large deletions.

When the triplet repeat tract within the plasmid was interrupted with another repeat motif the level of deletions observed varied. Interrupted repeat tracts included (GCT)27ACT(GCT)57, (GCT)25ACT(GCT) and (GCT)27ACT(GCT)20ACT(GCT)106 as well as a range of pure repeats (Parniewski *et al.*, 2000). It was observed that in a strain of *E. coli* with a functioning MMR that larger deletions occurred in disrupted tracts. It is possible that the transition between the repeat motifs stalls DNA polymerase function in a manner not observed in pure tracts allowing single strands more time to form alternative structures.

It has already been discussed that the mutation of mismatch repair gene products effect the stability of microsatellite DNA. In a study of HNPCC cases the majority of patients showing a germline mutation in a mismatch repair gene also showed microsatellite instability in their tumours (Liu, 1996). Cases of early breast cancer carcinoma showing microsatellite instability were examined across the trinucleotide repeat of SBMA, SCA type 1 and DM1. Approximately 5% of the tumours under examination exhibited triplet repeat instability, and when it was detected, it was present at all three loci. The DM1 locus was found to be the most unstable in all the tumours under investigation (Shaw *et al.*, 1996). The differences in the

stability levels at the different loci suggests there are *cis* acting factors effecting the degree of stability of the repeat DNA. For example quantification of repeat expandability at 10 unstable CAG·CTG loci has found that flanking sequence modifiers, in this case the GC content of the flanking region, influence the degree of expandability at these loci with identical repeat motifs (Brock *et al.*, 1999), while it has already been observed that the purity of the repeat tract effects its stability levels (Parniewski *et al.*, 2000).

1.10.2 *PMS2*

Although there are two *MutL* complexes, *MutLα* and *MutLβ*, it is the *MutLα* complex which has the primary role. This MMR protein complex consists of both *MLH1* and *PMS2* however, examination of the outcome of the loss of either protein on MMR finds them to have an unequal effect. For example, examination of the rate of mutation at long and short tracts of mononucleotide repeats in murine null homozygotes, *Mlh1* (-/-) or *Pms2* (-/-), observed that the mutation rate is 2-3 times higher in *Mlh1* (-/-) mice. Interestingly comparison of the mutator phenotype in *Mlh1* (-/-) mice with those deficient in *Mlh1* (-/-) and *Pms2* (-/-) found there to be little difference (Yao *et al.*, 1999). It was also noticed that the tumours developed by *Mlh1* (-/-) or *Pms2* (-/-) mice also differed. Yao *et al.* (1999) observed that 83 % of the 24 *Mlh1* (-/-) mice studied developed intestinal tumours, in contrast, although 95% of the 20 *Pms2* -/- mice examined developed tumours by 6 months they were mostly lymphomas and the remainder sarcomas (Qin *et al.*, 1999). These observations, coupled with the knowledge that loss of *MLH1* along with *MSH2* accounts for the majority of MMR mutations in HNPCC kindreds and the rarity of *PMS2* derived tumours (Prolla *et al.*, 1998) might suggest that *PMS2* has a residual role in MMR and its function in MMR is *Mlh1* dependent (Qin *et al.*, 1999). In fact, inactivation of *PMS2* results in different sorts of tumours to its *MutLα* partner may indicate differences in primary function of the proteins. Sterility in male *Pms2* (-/-) mice where problems were observed with chromosome alignment and pairing indicates a role in meiosis (Baker *et al.*, 1995), although sterility was observed in male and female *Mlh1* (-/-) mice (Baker *et al.*, 1996).

Investigation of the mutation frequency at the *hypoxanthine-guanine phosphoribosyltransferase* (*hprt*) locus in the human uterine tumour cell line, HEC-1-A, a cell line which produces a truncated human PMS2 protein product, observed the frequency of mutations to be around 100 times higher than those observed in cells proficient in wild type

32

repair and all of the observed mutations were point mutations. The majority of which 66% were +/- 1 bp insertion / deletions, occurred at repetitive mononucleotide regions (Kato *et al.*, 1998). This effect was observed to be in common with that observed at the *hprt* locus in MLH1 deficient tumour cell lines (Ohzeki *et al.*, 1997). However, C→T transitions at a CpG site in MLH1 deficient tumour cell lines were observed to be high, this was not the case in the HEC-1-A cell line. This reduction in C→T transition at CpG sites was in common with a similar observation made in *Pms2* (-/-) mice (Narayanan, 1997) suggesting that PMS2 may have a role in the repair of the mismatch that results in C→T transition at CpG sites.

1.11 Aims and General Strategy

The progressive expansion of the triplet repeat associated with *DMPK* is associated with increasing severity of symptoms observed in DM 1 patients. Although there are several hypothesis as to the mechanism of expansion there is little data available about the progression of repeat instability within either somatic tissues or germline transmissions.

Several lines have been generated using the *Dmt* transgene (Monckton *et al.*, 1997) (Figure 3.2) and initial study of somatic tissues of the *Dmt* lines has revealed that there is little instability in the seven tissues examined in the young mice of the *Dmt*-B, *Dmt*-C and *Dmt*-E lines. A small amount of repeat instability, +/- 5 repeats, was observed in the *Dmt*-D line. Maternal and paternal germline instability was however, observed in all four of the lines investigated.

The aim of this project was to characterise the *Dmt* transgenic mice further and examine the effect of *cis* and *trans* acting modifiers on the repeat. Specifically I have:

- Investigated the level of somatic instability in mature mice and the progression of any repeat instability.

- Investigated whether or not the transgenes are transcribed and the effect this may have on triplet repeat instability.

- Studied the repeat transmission pattern and how this develops over the lifetime of the transmitting parent.

- Investigated the effect a mutated mismatch repair mechanism has on the stability of the repeat.

34

Chapter 2

Materials and Methods

2.1 Materials

2.1.1   Chemicals and Reagents

Chemicals of molecular biology grade were obtained from Merck Ltd. (BDH Laboratory Supplies) and Sigma Chemical Company Ltd.. Reagents obtained from alternative sources are mentioned when appropriate.

2.1.2   DNA sources

2.1.2.1 *Dmt*-D 3111 kidney cell line

Dmt-D 3111 kidney cell line was provided by Mr Mario Pereira, Division of Molecular Genetics, University of Glasgow.

2.1.2.2 Mice

The *Dmt* transgenic mice were generated by Dr Darren Monckton (Monckton *et al.*, 1997) and I was given access to archived tissues as well as breeding mice from each line, *Dmt*-A, -B, -C, -D and -E. All *Dmt* mice were on an FVB/n background. Supplementary FVB/n wild type mice were obtained from Harlow, UK, Ltd.

C57/black 6 mice, transgenic for the disrupted murine *Pms2* gene (Baker *et al.*, 1995) were obtained and bred onto the FVB/n background.

2.1.2.3 DNA markers

All DNA samples that were electrophoresed on agarose gels were measured against *Hae* III digested ϕ X174 DNA and / or *Hin*d III digested λ DNA. Lambda and ϕ X 174 DNA were supplied by New England Biolabs.

When fluorescent labelled PCR products were sized on a 6% acrylamide denaturing gel using Genescan® software the size of the product was estimated against either ROX 1000 or ROX 2500 molecular weight size standard (Applied Biosystems, Perkin Elmer).

2.1.2.4 Probes

DNA probes used included, DM56, a CTG•CAG repeat PCR product, DM-F / DM-PRENK, a non repeat *Dmt* transgenic PCR product, a 1019 bp fragment from the mouse *Usf2* gene, amplified with mUSF-A and mUSF-BR, a *Dmpk* PCR product generated between exon 2 /exon 4 and PCR products generated with mPMS2-1, mPMS2-2 and mPMS2-3. The probes were generated using PCR, resolved on an agarose gel and purified using the Qiagen PCR Purification Kit.

2.1.3 Enzymes

Restriction enzymes and their buffers were obtained from Boehringer Mannheim (BM), Life Technologies and New England Biolabs. Other enzymes and their source are listed in table 2.1.

Enzymes	Company
FPLC pure klenow, cloned	Pharmacia
Amplitaq	Perkin Elmer
Biotaq	Bioline
RNasin	Promega
Proteinase K	Boehringer Mannheim
Superscript™ II	Gibco BRL
RQ1 RNase free DNase	Promega

Table 2.1 Enzymes and suppliers

2.1.4 Hormones

PMSG, and HCG, were both supplied by Sigma.

2.1.5 Kits

The kits listed in table 2.2 were used during the project.

Kit	Supplier
Bac II, soft tissue extraction kit	Nucleon
Qiagen Plasmid DNA Mini prep kit	Qiagen
Qiagen Gel Purification Kit	Qiagen
Ready-To-Go™ DNA Labelling Beads (-dCTP)	Pharmacia Biotech
TOPO TA Cloning®	Invitrogen®

Table 2.2 Kits and suppliers

2.1.6 Oligonucleotides

The oligonucleotide primers were supplied by Genosys. The fluorescent oligonucleotides were supplied by Oswell. The primer sequences are listed in table 2.3. The locations of *Dmt* transgenic primers are indicated in Figure 2.1.

nucleotide	sequence 5'-3'	TM	target sequence
CLW2F	GAATTCAGGCTTAAGGAGGTCCGACTG	61°C	*Dmpk*, exon 2
CLW4R	GAATTCGCAAAATGCAGCTGTGTGATC	61°C	*Dmpk*, exon 4
DM-BR	CGTGGAGGATGGAACACGGAC	68°C	*Dmt* transgene (9)
DM-C	AACGGGGCTCGAAGGGTCCT	66°C	*Dmt* transgene (4)
DM-DR	CAGGCCTGCAGTTTGCCCATC	68°C	*Dmt* transgene (7)
DM-DRF	DM-FAM-linker-CAGGCCTGCAGTTTGCCCATC	68°C	*Dmt* transgene
DM-DR12	CAGGCCTGCAGT	40°C	*Dmt* transgene (8)
DM-ER	AAATGGTCTGTGATCCCCCC	60°C	*Dmt* transgene (6)
DM-GR	GCAGGGCGTCATGCACAAGAAA	68°C	*Dmt* transgene (10)
DM-H	TCTCCGCCCAGCTCCAGTCC	74°C	*Dmt* transgene (2)
DM-H12	TCTCCGCCCAGC	38°C	*Dmt* transgene (3)
DM-PRENK	GTCCGGTACCGAATTCCGCTAGCTCCTCCCAGACCTTC		*Dmt* transgene (12)

nucleotide	sequence 5'-3'	TM	target sequence
DM-QR	CACTGTGGAGTCCAGAGCTTTG	66°C	*Dmt* transgene (11)
DM-R	GTCCTCCGACTCGCTGACAG	68°C	*Dmt* transgene (5)
mDmtD-I	CCAACGTTCATTTGCCATTTTTAGAC	60°C	*Dmt*-D novel sequence (1)
mGDH-A	CACTTGAAGGGTGGAGCCAAAC	69°C	mouse *Gapdh*
mGDH-BR	TGGGTGGTCCAGGGTTTCTTAC	68°C	mouse *Gapdh*
mP2-1	TTCGGTGACAGATTTGTAAATG	61°C	mouse *Pms2* (Baker *et al.*, 1995)
mP2-2	TTTACGGAGCCCTGGC	62°C	mouse *Pms2* deficient homozygote (Baker *et al.*, 1995)
mP2-3	TCACCATAAAAATAGTTTCCCG	61°C	mouse *Pms2* (Baker et al., 1995)
mUSF-A	GCCCCTGCCTCACCGTATAG	66°C	mouse upstream stimulatory factor 2
mUSF-BR	CTGGGGTCCACCACTTCAAG	67°C	mouse upstream stimulatory factor 2
XYT	CGAGGGGGATGGTCGACGGAAGC GACCTTTTTTTTTTTTTTTTTT		anneal to poly A sequence (Friedrich et al., 1993)

Table 2.3 Oligonucleotides
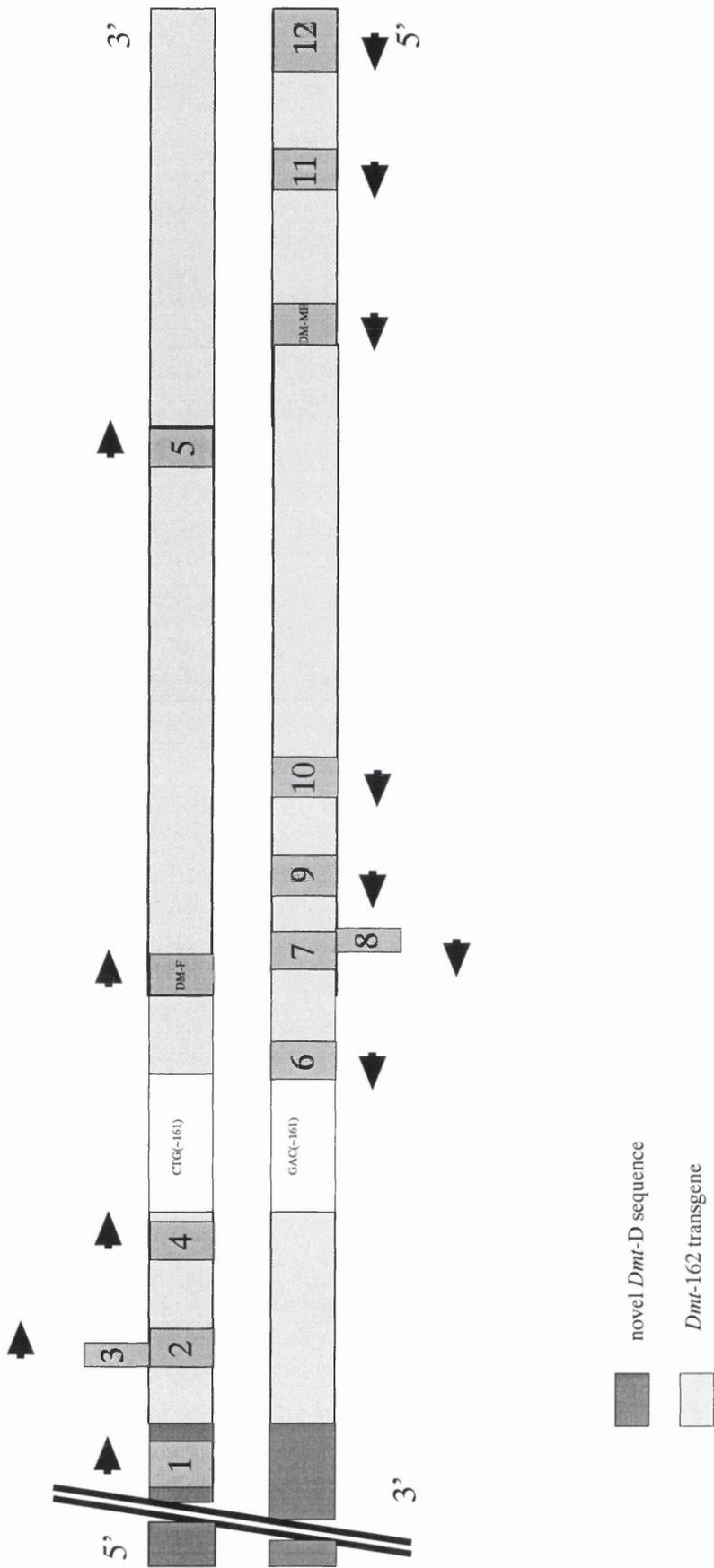
Figure 2.1
Location and orientation of primers on *Dmt* transgene
Schematic diagram of *Dmt* transgene and the location and orientation of the DM and mDmtD primers. The primer name and sequence is given in Table 2.3

novel *Dmt*-D sequence

Dmt-162 transgene

2.1.7 Solutions

Solutions were sterilised by autoclaving at 121 (C for 15 minutes. Filter sterilisation was carried out by passing the solution through a Gelman Science 0.2 (M pore filter.

Ampicillin
Stock solution: 50 mg ml^{-1}, store at –20 °C
Working solution: 50 µg/ ml^{-1}.

Cell lysis buffer
500mM Tris-HCl (pH 9.0), 50mM EDTA, 2.5% SDS (w/v).

Denaturing solution
1.5M NaCl, 0.5M NaOH.

Depurinating solution
0.25 M HCl.

DNA Dilution buffer
Stock solutions of 50 X Te and 10µM primer are diluted to a working stock of 1X Te and 0.1 µM oligonucleotide.

5 X 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®400, 3 X TBE.

DL-Dithiothreitol (DTT)
Stock solution, 0.1M in H$_2$0 (stored at -20°C).

11 X PCR buffer
At 1 X concentration in the final PCR amplification, 45 mM Tris-HCl (pH 8.8), 11mM ammonium sulphate, 113 µgml^{-1} BSA, 1mM dATP, 1mM dCTP, 1mM dGTP, 1mM dTTP, 4.4 µM EDTA, 4.5 mM MgCl$_2$, 6.7 mM 2-mercaptoethanol.

Low concentration dNTP 11 X PCR buffer

At 1 X concentration in the final PCR amplification, 45 mM Tris-HCl (pH 8.8), 11mM ammonium sulphate, 113 μgml^{-1} BSA, 0.2mM dATP, 0.2mM dCTP, 0.2mM dGTP, 0.2mM dTTP, 4.4 μM EDTA, 4.5 mM MgCl$_2$, 6.7 mM 2-mercaptoethanol.

Ethidium bromide

Stock solution: 10 mgml^{-1}.

Working solution: 200 ngml^{-1}.

Gelatine coated slides

Glass slides were coated in 0.1% gelatine (w/v) in 1X PBS and incubated for ~ 3 hours at 37 °C.

5 X Genescan blue loading dye

Ficoll 400-DL (50 mg ml^{-1}), Dextran sulphate (MW 500, 1.7 mg ml^{-1}), Blue dextran (41.5 mg ml$^{-1)}$ in 2X TBE. Store at 4 °C.

Genescan mix

1 X Genescan blue loading dye, 27% ROX 2500 or 1000 (v/v) and 63% deionised formamide (v/v).

Ipegal CA-630 lysis buffer

0.65 % Ipegal CA-630 (Sigma), 10 mM Tris-HCl pH 7.9, 150 mM NaCl, 1.5 mM MgCl$_2$.

Lysis buffer

50 mM Tris (pH8.0), 100 mM EDTA (pH8.0), 0.5% SDS (w/v), 600 μgμl^{-1} proteinase K.

Neutralising solution

1.5 M NaCl, 0.5M Tris H-Cl (pH 6.5).

Orange G

0.06 % (w/v) Orange G, 50 % (v/v) glycerol.

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.

10 X Phosphate buffered saline (PBS)

1.4 M NaCl, 0.027 M KCl, 0.1M Na_2HPO_4, 0.021 M KH_2PO_4 (pH 7.2 with HCl).

Phosphate hybridisation solution

Solution A: 14% w/v SDS, 2 mM EDTA; Solution B: 1M $NaPO_4$ (4°C). Was mixed A:B, 50:50, and warmed to 65 °C.

Potassium Hydroxide lysis buffer

200mM KOH.

Proteinase K

Working stock of 20 mgml^{-1} in filter sterilised H_2O.

Proteinase K Lysis buffer

50 mM Tris (pH8.0), 100 mM EDTA (pH8.0), 0.5% SDS, 600 $\mu g \mu l^{-1}$ proteinase K.

RNA loading buffer

10 mM NaH_2PO_4 (pH7.0), 50% (v/v) glycerol, 0.4% (w/v) bromophenol blue.

Sephadex G50 (Pharmacia Biotech)

Sephadex G50, equilibrated with TE (pH7.6).

10% (w/v) SDS

10% (w/v) SDS in H_2O.

20 X SSC

6 M NaCl, 0.6 M $Na_3C_6H_5O_7.2H_2O$.

20 X SSPE

0.2 M $NaHPO_4.2H_2O$, 3 M NaCl, 0.02 M EDTA (pH 8.0).

Standard cell culture medium

10% Fetal Bovine Serum, 100 Units ml^{-1} penicillin, 100 μg^{-1} streptomycin in Dulbecco's modified eagle medium, all supplied by Gibco BRL, Life technologies.

1 X TAE

0.04M Trizma base, 0.04M acetic acid, 0.001M EDTA (pH 8.0).

0.5 X TBE

0.045M Trizma base, 0.045 orthoboric acid, 0.001 M EDTA (pH 8.0).

Te

10 mM Tris-HCl (pH 8.0), 0.1 mM EDTA (pH 8.0), and autoclaved.

TE

10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0), and autoclaved.

Tricene buffer

200 mM Tricene in distilled water and autoclaved.

1M Tris (pH 8.0)

1M Trizma base and HCl to pH 8.0.

1M Tris (pH 7.9)

1M Trizma base and HCl to pH 7.9.

Tris-HCl (pH 9.0)

1M Trizma base and HCl to pH 9.0.

Oligo labelling buffer

Oligo labelling buffer was made up to the ratio 2 A: 5 B: 3:C and stored in aliquots at –20 °C.

Solution A: 1.2 M Tris (pH 8.0), 121 mM $MgCl_2$, 0.48 mM dATP, 0.48 mM dGTP, 0.48M dTTP, 18 µl of 2- β-mercaptoethanol in a final volume of 1033 µl.

Solution B: 2 M Hepes, pH6.6 with NaOH (4 °C).

Solution C: Pharmacia Biotech Ultrapure dNTPs set suspended in 0.2 mM EDTA (pH 7.0) to an Optical Density of 90 OD units ml^{-1} (-20 °C).

X-gal

50 mg ml^{-1} of 5-bromo-4-chloro-3-ondoyl-ß-D-galactoside (X-gal) in dimethylformamide. Store at -20 °C.

2.2 Methods

2.2.1 Polymerase chain reaction

All PCR amplifications were carried out in either a Biometra Uno thermal cycler or a Biometra T 3 thermal cycler, in either Costar 96 well plates or Anachem thin walled 200 µl tubes. Each reaction was overlaid with mineral oil (Sigma) and then sealed. The thermal cycler lid was preheated to 105 °C. Every reaction was carried out with either Promega Amplitaq or Bioline Biotaq.

2.2.1.1 Polymerase chain reaction conditions

Except where stated otherwise each PCR amplification was carried out in a volume of 7 µl with 0.175 units of *Taq* polymerase in which 11 X PCR buffer had been diluted to 1 X and each primer was at a final concentration of 1µM. The reactions were cycled under the following PCR conditions:

96 °C for 45 seconds,

*°C for 45 seconds,

70 °C for 3 minutes for 30 cycles

* °C for 1minute,

70 °C 10 minutes for 1 cycle, hold at 4 °C.

Where * is the primer annealing temperature.

2.2.2 Agarose gel electrophoresis

Gels of between 0.8 and 3 % were made up with Agarose MP, Boehringer Mannheim and / or Nusieve agarose (FMC BioProducts), with either 0.5 X TBE or 1 X TAE and cooled with agitation to ~ 50 °C before the gel was poured. Ethidium bromide (EtBr) was either added to the gel at a concentration of 200 ngml^{-1}, just before pouring, or used to stain the gel after it had been electrophoresed by shaking the gel for ~ 1 hour in distilled water containing EtBr.

Each gel was electrophoresed for between 1 and 16 hours and at between 20 and 300 V. At least 1 lane per gel was run with 1.0 µgµl^{-1} of λ *Hin*d III / φX174 *Hae* III marker. The DNA samples / marker were visualised using a UV transilluminator (λ 260 nm) and photographed onto thermal paper (Mitsubishi) using a UVP gel documentation 7500 system.

2.2.3 Maintenance of mouse colony

2.2.3.1 Mouse tailing

When each mouse was at least 4 weeks old, it was anaesthetised using Fluorothane, Zeneca Ltd., (halothane, 100% w/w) a numbered tag placed in its ear, ~ 1cm of tail was removed with a sterile scalpel and the tipped tail cauterised. The removed tail was placed in 500 µl of proteinase K lysis buffer and incubated overnight at 60 °C.

2.2.3.2 Mouse typing and sizing

After overnight incubation each tail was vortexed vigorously and debris pelleted. 1 µl of the lysate was removed and diluted 1/100 in distilled water, assuming 1 µl of diluted lysate to contain ~ 1 ng of DNA. The proteinase K was heat inactivated by incubating the dilution at 95 °C for 5 minutes. This diluted lysate was used as the template for PCR amplification to determine the presence of the transgene. The same diluted lysate was also used as the template for transgenic repeat sizing.

2.2.3.3 Phenol / phenol: chloroform extraction of DNA

To purify the DNA from each tissue an equal volume of phenol was added to the tissue lysate, mixed thoroughly by vortexing, and centrifuged at 10 000 rpm in a bench top centrifuge. The upper layer was removed and placed into a clean tube before adding an equal volume of phenol : chloroform : isoamyl alcohol, 25 : 24 : 1, vortexing and centrifuging. The upper layer was again removed and placed in a clean tube. The phenol : chloroform : isoamyl alcohol extraction was repeated if necessary. A volume of 2M sodium acetate 1/10 of the aqueous fraction was added and mixed. The DNA was precipitated by adding 7/10 volume of isopropanol, mixing and spinning for 5 minutes at 10 000 rpm in a bench top centrifuge. The pelleted DNA was washed in 1 ml of 80 % ethanol (v/v), the supernatant removed and the pellet air dried for ~ 20 minutes. The pellet was resuspended in an appropriate volume of Te, and incubated for 1 hour at 60 °C. To determine the concentration of the DNA it was diluted 1/ 80 in distilled water to a final volume of 400 µl and the absorbance of the DNA at λ 260 nm measured. The concentration of the DNA was determined by taking into account the dilution factor and that 50 mgml⁻¹ of DNA gives an absorbance of 1 at λ 260nm. Alternatively the DNA concentration was determined from an agarose gel. A known volume of DNA was electrophoresed on an agarose gel against a DNA marker of known concentration. A photograph was taken and scanned into Adobe Photoshop. The concentration of the unknown

DNA was determined from this image using the Kodak 1 dimension programme. All DNA was stored at -20 °C.

2.2.3.4 PCR product transfer by "squash blotting"

After the gel had been visualised using the transilluminator it was cut to an appropriate size including markers and rinsed in distilled water. The gel was then shaken at 100 rpm in a series of solutions, with thorough rinsing in distilled water between each stage. Firstly depurinating solution was added for not more than 10 minutes, followed by 30 minutes in denaturing solution and 30 minutes in neutralising solution. The gel was placed onto two pieces of Whatman 3 MM chromatography paper soaked in neutralising solution cut to the size of the gel. A piece of nylon transfer membrane pre soaked in distilled water followed by neutralising solution was placed on top of the gel and all bubbles and creases removed. Membrane for the transfer of DNA was either Hybond-N nylon DNA transfer membrane obtained from Amersham International PLC or Magna nylon membrane obtained from Micron Separations Inc through GSI. Two further layers of neutralising solution soaked Whatman were placed on top followed by paper towels, a glass plate and a weight. After transfer of between 3 and 16 hours the blot was disassembled and the membrane baked at 60 °C for 40 minutes before exposure for 30 seconds to UV. The filter could then be stored indefinitely between dry Whatman paper.

2.2.3.5 Southern hybridisation of squash blot transfer filters

Each filter was soaked in distilled water before being transferred into a hybridisation tube. The phosphate based hybridisation solution was prewarmed to 65 °C before adding 5 mls to each hybridisation tube and incubating the filter at 65 °C in a Biometra OV 5 hybridisation oven. The buffer was changed after 20 minutes and this was repeated once more before the appropriate probe was added. The filters were incubated with the probe overnight before rinsing with wash solution, twice in the hybridisation oven and once flat in a shaking tray. Each wash was for 20 minutes. The membranes were then exposed to X-ray film supplied by Konica and developed using a X-Ograph Compact X2, (X-Ograph Ltd).

2.2.4 Labelling probes for Southern hybridisation

2.2.4.1 Labelling probes for Southern hybridisation with oligo labelling buffer

40 ng of probe of interest and 2 ng of λ Hind III / φX174 Hae III marker were made up to 20 μl in distilled water boiled for 2 minutes and cooled on ice. 6 μl of OLB, 1.2 μl of BSA, and 25 μCi of [αP ^{32}] dCTP, (3000 Ci/nmol), (ICN or Amersham) and 1 unit of Klenow were then added before incubating at 37 °C for at least 1 hour.

2.2.4.2 Labelling probes for Southern hybridisation with Pharmacia Biotech Ready-To-Go™ DNA Labelling Beads (-dCTP)

20 ng of probe of interest and 2 ng of λ Hind III / φX174 Hae III marker were made up to 20 μl in distilled water, boiled for 2 minutes and cooled on ice. A Pharmacia Biotech Ready-To-Go™ DNA Labelling Bead (-dCTP) was resuspended in 10 μl of distilled water before the precooled probe was added. To this mixture 50 μCi of [αP ^{32}]dCTP was added, mixed and the reaction incubated at 37 °C for at least 30 minutes for a triplet repeat template and less than 30 minutes for a simple template.

2.2.4.3 Clean up of labelled oligo by a Sephadex® G50 spin column

The reaction was made up to 200 μl with distilled water before passing through a Sephadex ® G50 spin column. The probe was boiled for 5 minutes before adding to the prehybridised filters.

2.2.4.4 Clean up of labelled probe by precipitation

Unincorporated radioactivity and dNTP's were removed from the labelled probe by precipitation. 30 μl of 2M Na Ac, 90 mg of salmon sperm DNA, (Sigma) and 140 μl of distilled water were added and mixed before adding 500 μl of 100% ethanol. The DNA was pelleted by briefly centrifuging in a bench top microfuge. The supernatant was removed and the pellet washed with 1 ml of 80% (v/v) ethanol. Again, the supernatant was removed and the pellet resuspended in 150 μl of distilled water. The probe boiled for 5 minutes before adding to the prehybridised filters.

2.2.5 Mouse dissection

2.2.5.1 Removal of blood from mouse

Immediately after sacrifice, a 23 gauge needle was placed into the thoracic cavity and blood removed into a syringe. This blood was quickly transferred into a 1.5ml tube containing 500 μl

1X SCC, 10 mM EDTA and mixed thoroughly to prevent clotting. The blood was frozen overnight, thawed and the white cell nuclei pelleted by centrifugation. The supernatant was decanted off and the white cell nuclei resuspended and washed in 1 ml of 1 X SSC. Resuspension of the pellet was achieved by vortexing. The rinsed nuclei were again pelleted and the supernatant removed. The pellet was resuspended in 100 µl of proteinase K lysis buffer and was incubated overnight at 60 °C. The DNA was then extracted by the phenol / phenol chloroform method.

2.2.5.2 Treatment of tissues removed from mouse

The tail, muscle, testes / ovary, uterus, colon, kidney, spleen, liver, diaphragm, lung, heart, tongue, brain and eyes, were dissected out, diced where necessary and placed in 500 µl of proteinase K lysis buffer. "Dirty" tissues such as the colon were "washed" in 1 X SSC and any "matter" removed before lysis. Dilutions were made from each proteinase K tissue lysate, assuming that 1 µl of the lysate contained ~ 1 ng of DNA. Serial dilutions of 1/10 (~ 1000 pgµl^{-1}), 1/100 (~ 100 pgµl^{-1}), 1/1000 (~ 10 pgµl^{-1}) were carried out in filter sterilised water or dilution buffer containing 0.1µM DM-H. 1µl of the 1/10 and 1/1000 serial dilutions was used as templates for small pool PCR. Alternatively DNA was purified from 250 µl of the proteinase K lysate and the concentration estimated by spectrophotometry. The DNA was diluted to 1200 pgµl^{-1}, 120 pgµl^{-1}, 12 pgµl^{-1} in dilution buffer containing 0.1µM DM-H. The 1200 pgµl^{-1} and 12 pgµl^{-1} concentrations were used as templates for tissue comparison using small pool PCR. A master mix containing everything except the appropriate DNA was dispensed into a 96 well plate. The DNA under investigation was added to the mastermix at an appropriate concentration. A small pool PCR amplification was carried out in duplicate for each tissue at each concentration for each mouse. The amplifications for each mouse at a specific concentration were resolved on the same gel, this made the detection of subtle differences in stability possible. Water was added in duplicate to the stock master mix to identify external contaminants.

2.2.5.3. Preparation of sperm

The vas deferens was removed and placed in a petri dish and covered in several microlitres of 1 X SSC. One end was held firm using forceps and using curved forceps sperm squeezed out of the vas deferens. Alternatively, the plugged uterus of a wild type mouse, set up in a mating with the male mouse of interest was removed after sacrifice and placed into a petri dish

containing 1 X SSC. The sperm was flushed out using 1 X SSC. The sperm / 1X SSC suspension was placed into a fresh microfuge tube. DNA was prepared from the sperm using the Nucleon Bac II kit and a protocol scaled down for a starting volume of 200 μl of sperm suspension.

2.2.5.4 Super ovulation of mice for the isolation of single oocytes

All of the injections and isolation of the ova were carried out by Mr Tony McDermott of the University of Glasgow Biological Services. Each injection was carried out intraperitoneally. Day one, each female mouse was injected with 5 units of PMSG.

Day three, (+47 hours after injection 1) 5 units of HCG were injected and the mouse mated with a vasectomised male overnight.

Day four, plugged females were sacrificed the oviduct removed and individual ova flushed out into M2 media.

2.2.6 Preparation of DNA for small pool PCR analysis

Small Pool PCR is a sensitive technique which is based on the principle of diluting known concentrations of DNA into "small pools" to be used as template for PCR. Thus small numbers or known numbers of molecules can be amplified and visualised. Its sensitivity makes it vulnerable to contamination and all stages of small pool PCR were carried out in a laminar flow hood, using filter tips and dedicated equipment. The lower concentration dilutions are diluted as appropriate on the day of amplification to ensure amplification of the required concentration.

2.2.6.1 Preparation of DNA for small pool PCR analysis by extraction with phenol

After overnight incubation in proteinase K lysis buffer the lysed tissue was vortexed vigorously before debris was pelleted. 250 μl of lysate was removed and DNA extracted by phenol, phenol / chloroform extraction. 1μg of DNA was digested for 1 hour with 1 unit of *Hin*d III, 1mM spermidine in a final volume of 200 μl, the reaction was incubated at 37 °C. The DNA was diluted to 1200 pgμl^{-1}, 120 pgμl^{-1}, 12 pgμl^{-1} in dilution buffer containing 0.1μM DM-H.

2.2.7 Preparation and manipulation of RNA

To determine whether the *Dmt* transgenes were transcribed, RNA was isolated from several tissues from each line. RNA is vulnerable to degradation from nucleases, to optimise the successful handling of RNA, the following precautions were taken. The restriction of RNA work to a dedicated bench, the use of dedicated equipment and stock reagents wherever possible and the wearing gloves at all times. All solutions required for RNA related techniques were made using DEPC treated water. DEPC was added to a concentration 0.1 % (v/v) to distilled water, mixed and allowed to stand overnight before the inactivation of the DEPC by autoclaving. All plasticware was soaked in active DEPC water before autoclaving, all glassware soaked in active DEPC water before baking. Each tissue was snap frozen in N (l) immediately after sacrifice and stored at -70 °C to minimise the action of RNases.

2.2.7.1 Extraction of RNA

When preparing RNA, tissue of known mass was defrosted on ice in a bijou before the addition of a minimum of 1ml of TRI REAGENT™(Sigma) per mg. The tissue was homogenised using a sterile 0.5mm Polytron probe with a Polytron Kinematica AG PT-3000 homogeniser and the RNA prepared according to the TRI REAGENT™ protocol. The pelleted RNA was dissolved in 75 µl of DEPC treated water and incubated at 65°C for ~ 20 minutes. To this 15 µl of RQ1, RNase free DNase (Promega, 1 unit µl⁻¹) and 10 µl of 10 X buffer was added. This was incubated at 37 °C for 30 minutes before being made up to 400 µl with DEPC water and extracted with an equal volume of phenol: chloroform: isoamyl alcohol (25: 24: 1) vortexed and centrifuged at maximum speed for 5 minutes. The upper layer was removed and precipitated with 40 µl of 2M sodium acetate and 280 µl isopropanol and precipitated at 13 000 rpm in a bench top centrifuge, for 5 minutes. The pellet was rinsed in 80% ethanol and air dried for not more than 30 minutes before being resuspended in DEPC water. The RNA concentration was determined by spectrophotometer by diluting 1/ 80 in distilled water to a final volume of 400 µl where the absorbance of the RNA at λ 260 nm was measured. The concentration of the RNA was determined by taking into account the dilution factor and that 40 mgml⁻¹ of RNA gives an absorbance of 1 at λ 260nm.

2.2.7.2 First Strand cDNA synthesis

Between 0.5 – 1 mM of oligonucleotides were added to 2.5 µg of total RNA and made up to a final volume of 9.5 µl with DEPC water for the reverse transcriptase positive (RT+) reaction. A no reverse transcriptase (RT-) control was made up to a final volume of 11.5 µl. The tubes

52

were then heated at 70 °C for 10 minutes, centrifuged briefly and placed on ice. 20 units of RNAsin (Promega), 10 µM of DTT, 1µM of dNTP's and Superscript II 5 X buffer were then added, to a final volume of 18 µl (RT+), 20 µl for the (RT-). Each sample was mixed and incubated at 42 °C for 2 minutes before 2 µl, 400 units of Superscript II reverse transcriptase was added to the (RT+) tubes only. The tubes were incubated 42 °C for a further 90 minutes. The reaction was heat inactivated by heating at 70 °C for 15 minutes. 10 µl of DEPC water was added to each reaction before storing the cDNA at - 70 °C.

2.2.7.3 Separation of cytoplasmic and nucleic cell fractions

This protocol was modified from Hamshere *et al.*, (1997), 3×10^6 cells from the *Dmt*-D 3111 kidney cell line, were washed in 1 X PBS and transferred into a 1.5 ml tube before the cells were pelleted in a bench top centrifuge. The supernatant was removed and the cells resuspended in 1ml of cell lysis buffer. The nuclei of the lysed cells were pelleted in a bench top centrifuge and 800 µl of the supernatant removed. This cytoplasmic fraction was transferred into a fresh tube. The remainder of the supernatant was removed from the pelleted nuclei and discarded. The nuclear fraction was resuspended in 400 µl of water. 100 µl of Ipegal CA-630 lysis buffer was added to the nuclear fraction and 200 µl to the nuclear fraction. RNA was then extracted from these fractions following the TRI REAGENT™ protocol and DNase treated.

2.2.7 Cloning and sequencing of PCR products

PCR products of interest were gel purified using the Qiagen PCR Purification Kit and cloned into the pCR®2.1-TOPO vector using the TOPO TA Cloning kit™ as described in the kit's protocol. The vector was transformed into F'*Escherichia coli* TOP 10 bacterial strain, with the genotype *mcr*A, Δ *(mrr-hsd*RMS- *mcr*BC), Φ80*lac* ZΔM15 Δ*lac* X74 *rec*A1 *deo*R *ara* D139 Δ *(ara, leu)* 7697 gal U galK rpsL (Str [R]) *end*A1 *nup*G.

Between 50-100 µl of transformed cell / SOC medium (Invitrogen®) mix was spread onto LB plates containing 50 µg/ml ampicillin and previously spread with 40 mg/ml of X-gal. The resulting white colonies were grown up overnight in LB medium containing 50 µg/ml ampicillin and purified using the Qiagen Plasmid DNA Mini Prep Kit.

The Molecular Biology Support Unit performed DNA sequencing on either an ABI 373A sequencer or an ABI 377 sequencer, at the University of Glasgow. Reactions were carried out as stated in the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit protocol.

Chapter 3

Somatic instability in the *Dmt* transgenic lines

3.1 Introduction

The expanded triplet repeat of DM1 patients accumulates somatic mosaicism over time (Wong *et al.*, 1995), the observed large expansions from the progenitor allele are apparently the consequence of multiple mutations over time giving rise to small length changes (Monckton *et al.*, 1995) the majority of which are expansion biased (Martorell *et al.*, 1998).

The mechanism of expansion is not known. It has been surmised that as a cell replicates the repeat tract could form alternative structures or "loop outs" which depending on the strand and the interpretation of the alternative structure by the mismatch repair complex might lead to expansions or deletions in the repeat tract (Figure 3.1). However, in older individuals it has been observed that the most severely affected tissue, the skeletal muscle, has acquired more increases in repeat length than that observed in blood (leukocytes) (Anvret *et al.*, 1993; Ashizawa *et al.*, 1993). This confounds the notion that the degree to which a cell replicates is the major mechanism of repeat expansion as the analysed blood DNA is derived from leukocytes, a cell population which is completely renewed over a matter of months. Cells are added to the skeletal muscle at a much slower rate, as the muscle develops and grows.

Our knowledge about the degree to which different tissues are unstable and how their instability levels relate to each other are limited. However, there appears to be a consensus that the smallest level of instability is observed in blood, with heart and kidney as well as muscle displaying the largest expansions (Jansen *et al.*, 1994; Joseph *et al.*, 1997; Lavedan *et al.*, 1993a; Tachi *et al.*, 1995a; Tachi *et al.*, 1993; Wohrle *et al.*, 1995; Wong *et al.*, 1995). Furthermore the limited data that is available about the degree of instability has been studied in a variety of ways. In the most common, digested genomic DNA is resolved on an agarose gel and transferred to a nylon membrane before the repeat was detected with suitable probe by Southern hybridisation, repeat instability was indicated by the diffuse nature of the hybridised band of the repeat DNA (Jansen *et al.*, 1994; Joseph *et al.*, 1997; Lavedan *et al.*, 1993a; Tachi *et al.*, 1995a; Tachi *et al.*, 1993; Wohrle *et al.*, 1995), although use of a more sensitive method of instability detection, small pool PCR is becoming more common
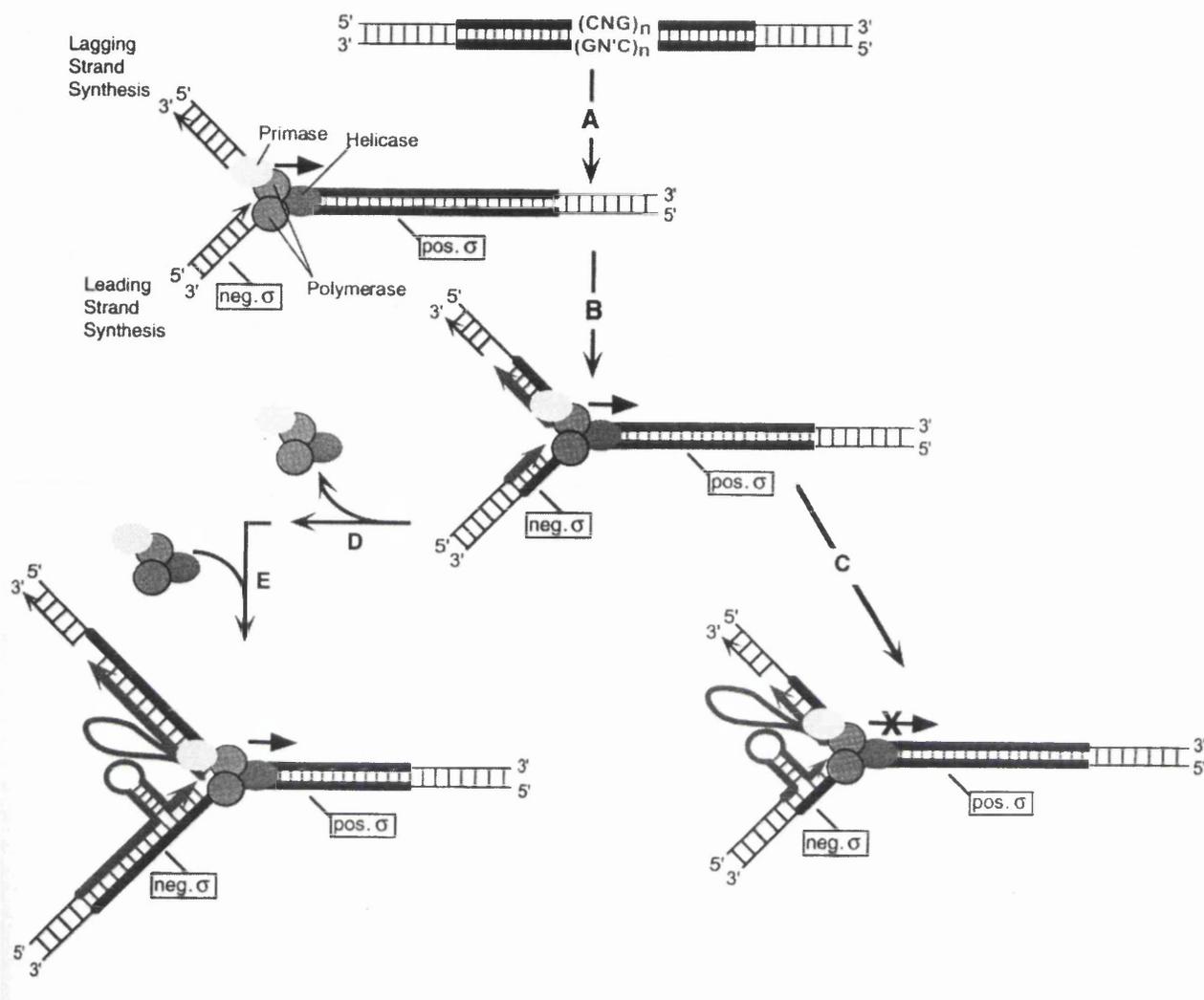
Figure 3.1.A
Model for triplet repeat expansion
Thick solid lines indicate the repeat regions, thinner lines correspond to the random sequence flanking DNA. N represents nucleotide T or G and N' represent nucleotide A or C. Pos and neg. refer to positive and negative supercoiling respectively. Figure taken from Gellibolian *et al.* (19

A

MutS beta — MSH3 | MSH2

MutL alpha — PMS2 | MLH1

B

repeat tract

paired DNA strand

template strand

replicating strand "stalls" in its passage through repeat tract

as thereplicating strand "stalls" or "slips" a "loop out" forms in the repeat tract

mismatch repair complex recognises the "loop out" and "repairs" it

resulting in an expansion

no change

resulting in a deletion

Figure 3.1.B
The replication slippage model of repeat instability
A. Representations of mismatch repair heterodimers, MutS beta and MutL alpha.
B. A figure demonstrating how the slippage of the replicating DNA strand as it passes through an expanded repeat tract could result in a change in the repeat length, and the role that the mismatch repair complex could play in the process.

56

(Martorell *et al.*, 1998; Wong *et al.*, 1995), as is displaying instability profiles using Genescan software® (Mangiarini *et al.*, 1997; Sato *et al.*, 1999; Wheeler *et al.*, 1999).

The study of mosaicism within DM1 is further confounded by the limits in the amount and nature of the tissues available, as well as the degree to which the disease has progressed within an individual. When a mature person is diagnosed after the onset of DM1 the progenitor repeat size of the disease range allele can only be surmised. Alternatively, might there be a mechanism affecting stability that is present only in certain tissue types? If multiple tissues were available for analysis might a pattern emerge which correlated with the affected tissues or that might suggest part of the instability process? Could the detection of such a pattern be restricted both by the limited range of tissues available, but also the diverse genetic backgrounds that the patients present? The advantage of a murine model for instability would be access to unlimited tissue of known progenitor size on a uniform genetic background.

The *Dmt* transgenic lines were generated to contain the *Dmt* 162 transgene (Figure 3.2) to provide a murine model for expanded triplet repeats which would allow investigation of the mechanisms of repeat instability. The repeat tract is asymmetrically flanked by only a small amount of non-coding DNA, 133bp on the 5' side and 609 bp on the 3' side. It contains no known promoter elements. This would suggest that the flanking transgenic DNA would have little effect on the repeat instability. Therefore the integration site of the transgene could instead account for any variation in stability levels between the almost identical transgenic lines. The *Dmt* mice provide access to all tissue types and a record of the progenitor repeat allele size is available. The size of the repeat of a tail sample taken from each mouse shortly after weaning at ~ 6 weeks was taken to be the progenitor repeat size. The investigation in the first instance was the level of somatic mosaicism in the *Dmt* lines. Of the five lines generated, *Dmt*-A was found to contain multiple copies of the transgenes and analysis of the somatic mosaicism in this line would be confusing. The other four lines *Dmt*-B, *Dmt*-C, *Dmt*-D and *Dmt*-E (Figure 3.2) are the result of simpler integrations and have been investigated. An initial study of somatic instability in young *Dmt* mice revealed little instability in most of the seven tissues examined. *Dmt*-D did exhibit a little expansion biased instability, although this was not greater than +/- 5 repeats (Monckton *et al.*, 1997).
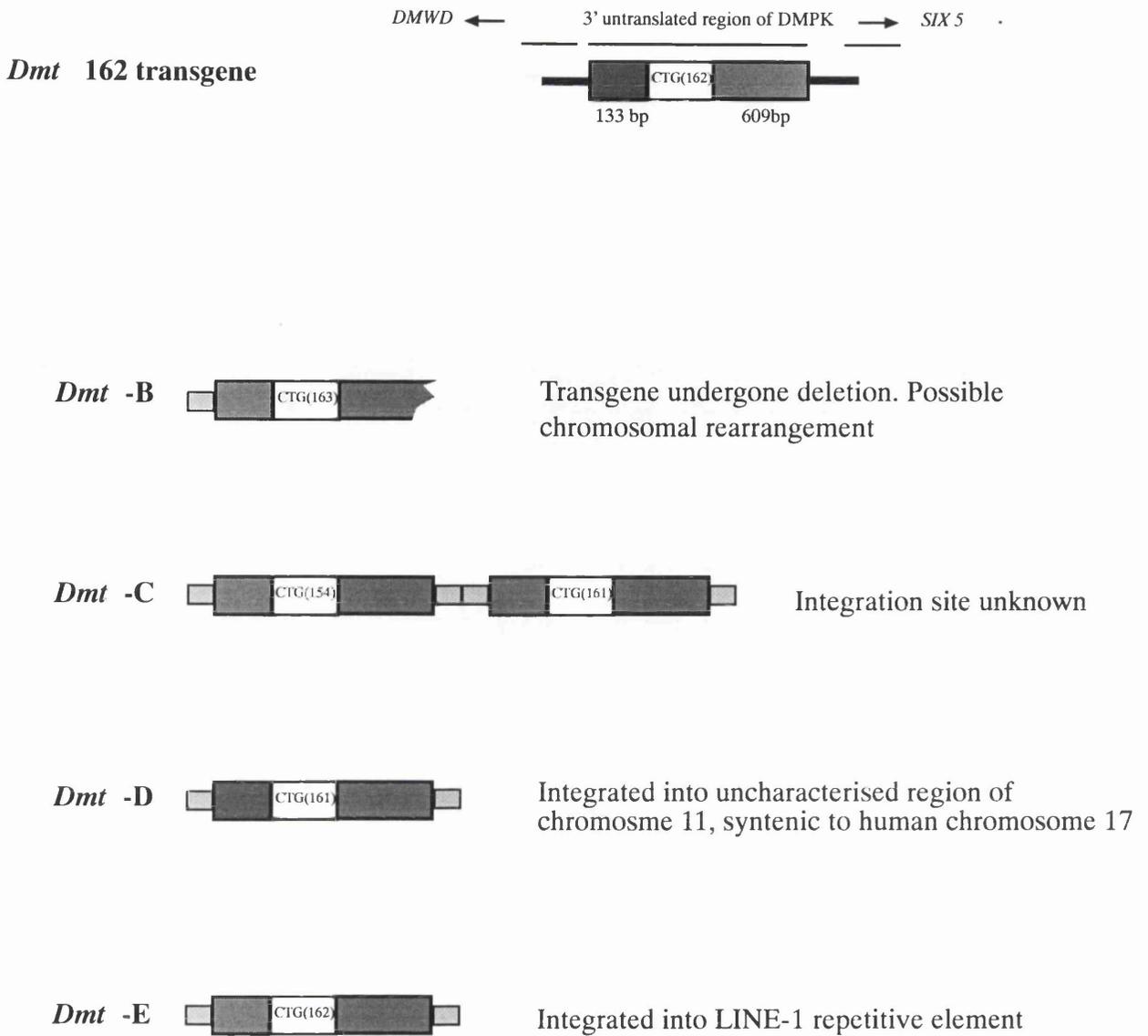
Figure 3.2 *Dmt* transgene and integrants (Monckton et al 1997)
Schematic representation of the *Dmt* transgene and the resulting integrants. The repeat size indicated for each integrant represents the repeat size of the founder mouse for each line

When the work encompassed in this chapter began little was known about the integration sites of the *Dmt* lines (Figure 3.2). An investigation into the integration sites of *Dmt*-D and *Dmt*-E was carried out by Dr Graham Brock of the University of Glasgow in parallel with my work. *Dmt*-D was found to have integrated into an uncharacterised region of chromosome 11, syntenic to human chromosome 17, while *Dmt*-E has integrated into a LINE-1 repetitive element. The aim of this chapter was to determine whether a murine model, with its short lifespan could be used:

1   to study a disease with a progressive component;

2   to establish the level of repeat stability in mature mice across a range of tissues in each of the lines; and

3   to establish whether the *Dmt* lines would be a suitable model with which to study the mechanisms and patterns of repeat expansion.

3.2 Identification of transgenic mice

The mice were bred on a FVB/n background. Each mating was set up between a parent hemizygous for the transgene *Dmt* (+/-) while the other was wild type (-/-). This meant that all of the mice examined in this part of the study were hemizygous for the transgene. The transgenic mice were identified from each litter by using transgene specific primers to amplify transgenic products. A 7 µl PCR reaction containing 1 µl of tail dilution ~ 1 ng of DNA, 1 X PCR buffer, 0.35 units of *Taq* polymerase and 1 µM of mouse typing mix 1 oligonucleotide primers (MTM-1) was set up. MTM-1 consists of the following primer pairs, mUSF-A and mUSF-BR which amplify a 1019 bp fragment from the mouse *Usf2* gene, DM-R and DM-QR which amplify a 175 bp none repetitive region of the transgene and DM-C and DM-ER which amplify across the transgenic repeat. The mouse *Usf2* gene PCR product was expected to be present in every successful PCR amplification irrespective of the presence of the transgene. Any MTM-1 reactions that did not result in a *Usf2* product were repeated. The presence of the non repetitive transgenic product and not the product including the repeat might indicate that the repeat was too large for the *Taq* polymerase to amplify. No case of this was observed during this study. Each primer in the stock MTM-1 was at a concentration of 10 µM. The transgenic typing PCR amplifications were cycled 30 times under the standard conditions with an annealing temperature of 63 °C. The samples were then electrophoresed on a 1.4 % agarose gel (w/v) in 0.5 X TBE. Photographs were taken (Figure 3.3.A) and the PCR products transferred by squash blot to a nylon membrane before hybridisation with a CTG•CAG repeat probe, DM 56 (Figure 3.3.B).

Figure 3.3
Identification of *Dmt* transgenic mice
PCR amplification to detect presence of the transgene by amplifying 1μl of tail lysate (~ 1 ng of DNA) in a 7μl volume containing 1μM of MTM primer mix in the final volume. MTM primer mix consists of mUSF-A and mUSF-BR, which amplify a 1019 bp fragment in all mice, DM-R and DM-QR which amplify a non-repeat region of the transgene and DM-C and DM-ER which amplify across the transgenic repeat, this will result in a product around ~515bp, but this band will vary in size according to the repeat that the mouse carries.
A: MTM PCR products were resolved on a 1.4 % agarose gel before being B: transferred onto a nylon membrane and hybridised with DM56. This CAG•CTG repeat probe will hybridise with the DM-C / DM-ER product only.

3.3 Examination of triplet repeat instability using small pool PCR

Small pool PCR is a technique whereby template DNA is diluted to known concentrations. This can result in the DNA template containing as little as an average of one molecule. This technique is therefore very sensitive as it can allow the amplification of individual molecules and an accurate estimation of the size of that molecule. When this is applied to microsatellite regions, the level of instability of a known allele can be monitored. Each SP-PCR amplification was cycled in a 7 μl reaction containing 1 X buffer, and 0.2 μM each of the forward and reverse primer, 0. 175 units of *Taq* polymerase and between 5 and 1000 pg of template DNA. The samples were cycled 28 times under standard PCR conditions with a suitable annealing temperature. Each SP-PCR amplification was made up to 10 μl with DNA loading dye before 5 μl of this sample was resolved on an agarose gel.

3.3.1 Optimal conditions for small pool PCR agarose gel electrophoresis

To ensure comparison of repeat alleles that may vary by as little as +/- 1 or 2 repeats could be achieved the PCR products were resolved on an agarose gel 40 cm in length. Low concentrations of template DNA means that the PCR product of a small pool PCR amplification was routinely not visible on the agarose gel. Therefore to visualise each small pool PCR amplification the DNA products were transferred onto a nylon membrane using the squash blot method of transfer and hybridised using the triplet repeat probe DM56. The repeat size was routinely estimated for each distinct molecule by sizing against the DNA marker. As estimation of the repeat size was reliant on this comparison the gel running conditions were optimised to ensure that each of the small pool PCR reactions and the markers progressed through the agarose gel at a steady rate relative to each other. Initially SP-PCR products were resolved on 1.5% agarose gels (w/v) in 0.5 X TBE at room temperature at 200V for 5 minutes followed by 130V O/N (~ 16 hours) at room temperature. These gels were found to "smile", that is the samples and markers did not resolve in a uniform manner (Figure 3.4.A). This was possibly due to the gel overheating during electrophoresis. All further SP-PCR electrophoresis was carried out at 4 °C and this improved the quality of the gels (Figure 3.4.B). Further improvements in the resolution of SP-PCR products were achieved by reducing the agarose gel percentage to 1.25% (w/v) in 0.5 X TAE and equilibrating the gel in 0.5 X TAE for ~1 hour before loading the samples (Figure 3.4.C, D and E). The optimal gel running conditions were determined to be a 1.25 % agarose gel (w/v), initial voltage was 300V for 15 minutes followed by 185 V for ~16 hours at 4 °C.

A

expected bands

mosaic bands

spurious bands

10 lanes: 1200 pg/µl   11 lanes: 120 pg/µl   11 lanes: 12 pg/µl

−272
−213
−161
−123

B

expected bands

spurious bands

−272
−213
−160
−123

C

175 −

−398
−298
−229
−139
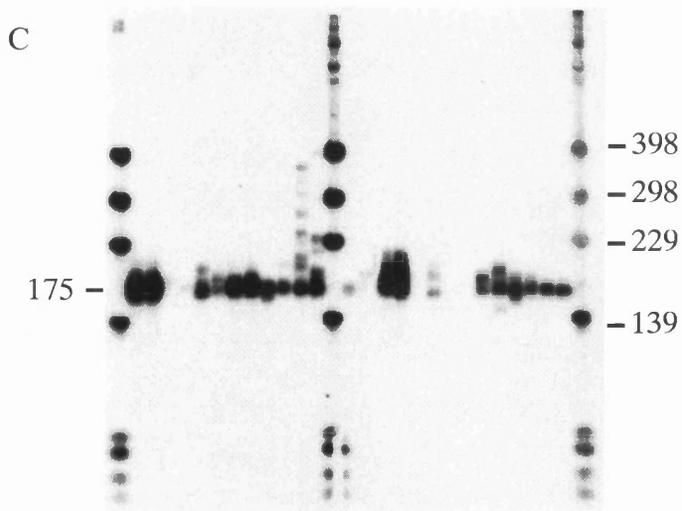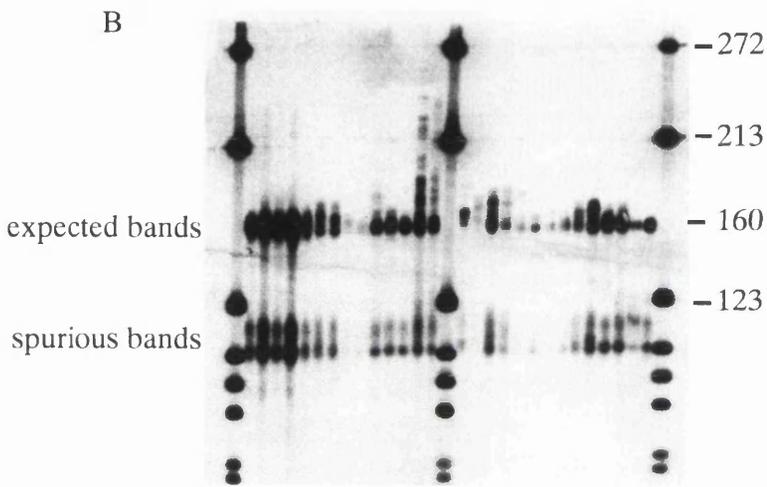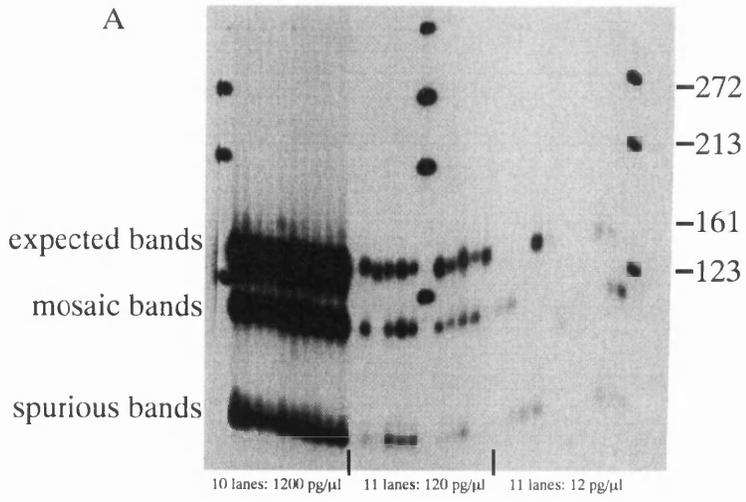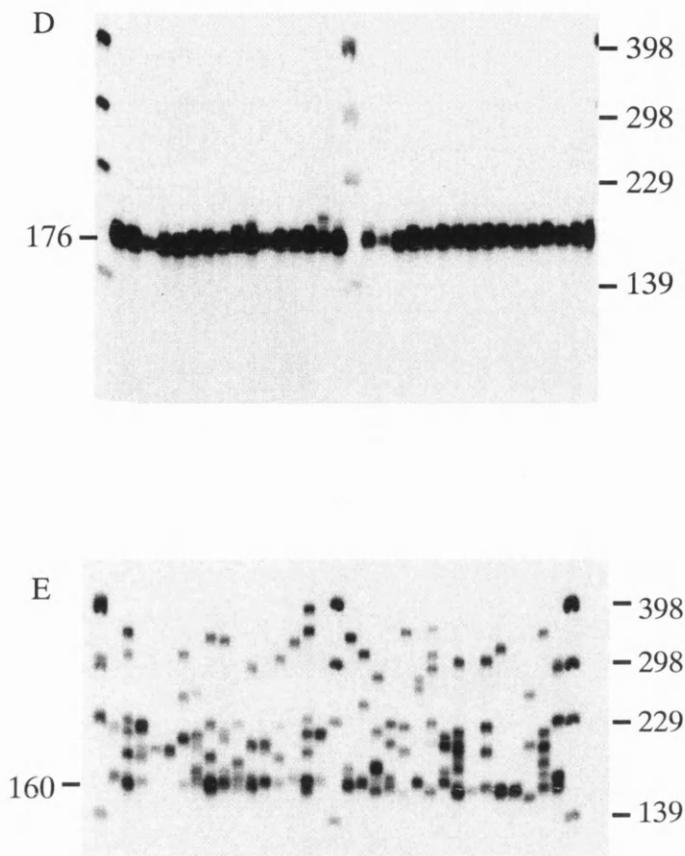
Figure 3.4 Small pool PCR products resolved using a selection of conditions

A The tail DNA of a *Dmt*-D mouse, mosaic for the transgenewas amplified using DM-H /DM-BR.This primer pair combination resulted in a secondary PCR product, ~ 90 repeats, which appeared to shadow the 161 repeat product pattern. It exhibits 2 transgenic repeat sizes of 161 and 110 repeats. It was assumed that this mouse was mosaic for both repeat sizes and that this mosaicism occured during embryonic development. Only the 161 repeat was transmitted to the offspring. The template DNA is at 3 different concentrations. The SP-PCR amplifications were resolved on a 1.5 % gel, at 200V for ~16 hours at room temperature.

B. Duplicate SP-PCR amplifications were carried out on a selection of tissues from a 13 month old *Dmt*-D mouse with161 repeats using DM-H /DM-BR. Approximately 1000pg of DNA was present in each reaction. The tissues from left to right are: spleen, blood, sk. muscle, lung, heart, kidney, marker, brain, colon, eye, diaphragm, tongue, mature tail and weaned tail.

The spurious secondary product can be seen as well as expected product. The SP-PCR amplifications were resolved on a 1.5 % gel, at 300V for ~16 hours at 4°C. The tissues are in the same order as B.

C. Duplicate SP-PCR amplifications were carried out on a selection of tissues from a 6 month old *Dmt*-D mouse using DM-H /DM-DR. Approximately 1000pg of DNA was present in each reaction. The SP-PCR amplifications were resolved on a 1.25 % gel, at 130V for ~16 hours at 4°C.

D. SP-PCR was carried out on ~100 molecules of lung DNA from a male mouse with 176 repeats. The SP-PCR amplifications were resolved on a 1.25 % gel, at 130V for ~16 hours 4°C.

E.SP-PCR was carried out on ~10 molecules of kidney DNA from a male mouse with 160 repeats. The SP-PCR amplifications were resolved on a 1.25 % gel, at 130V for ~16 hours at 4°C.

3.3.2 Choice and primer pair for small pool PCR amplification

The sensitivity of the small pool PCR technique for amplification from a template which may contain only 1 cells worth of DNA is reliant on the oligonucleotide primer pair chosen to amplify across the transgenic repeat resulting in a single band. When DM-H and DM-BR were used to amplify SP-PCR products it resulted in multiple PCR bands. The hybridisation of these bands with the DM56 probe confirms that these bands are transgenic in origin. It became apparent that one molecule could result in more than one transgenic band. As we wished to quantify the level of repeat stability this result would distort the resulting data. However, observation of the bands suggested a degree of shadowing occurring and suggested that one template molecule was producing a product of the expected size and also a smaller secondary (possibly single stranded) product (Figure 3.4.A and B). Whilst the template DNA was stable this shadowing effect had little influence on the interpretation of the results. However, when the molecules under examination are exhibiting instability there is a possibility that the true repeat product and its unstable shadow might confound the results (Figure 3.4.B). Alternative transgenic repeat spanning combinations were tried to determine the optimal SP-PCR combination. DM-H and DM-DR were found to result in a single product band for each molecule amplified (Figure 3.4.C, D and E).

It was suggested that the stringency of the small pool PCR reaction would be improved if the transgenic primers were further away from the repeat. In the case of *Dmt*-D, novel sequence was available on the 5' side (133bp) of the *Dmt* transgene (Figure 2.1). The efficiency of a primer, mDmtD-I, using this novel sequence was also examined (Figure 3.5).

Each PCR reaction was carried out in a 7 µl volume that contained, ~ 1ng of DNA, 1 X buffer, the forward and reverse primers under test at a 1µM concentration and 0.175 units of *Taq*. The PCR amplifications were cycled 30 times under standard conditions using the stated temperature. The PCR products were resolved on a 1.5% agarose gel (w/v) and photographed. The mDmtD-I and DM-BR primer combination produced no bands at all at annealing temperatures of 60, 62 and 64°C when mDmtD-I was combined with either DM-ER, DM-GR or DM-QR multiple bands occurred at 60, 62 and 64°C and increasing the annealing temperature to 66 and 68°C resulted in the loss of all bands. The combination of mDmtD-I with DM-DR also resulted in multiple bands at 60, 62 and 64°C as well as 66 and 68°C. Increasing the annealing temperature to 70, 72 and 74°C resulted in the loss of all bands. It should be noted that although not all of the bands would result in multiple misleading PCR products of transgenic origin, they would decrease the efficiency of the transgenic product reaction which can rely on template DNA from a single molecule. It was concluded that the most efficient primer pair to use to amplify across the transgenic repeat would be DM-H and DR at 68 °C. An advantage to this primer combinations is that their transgenic nature meant
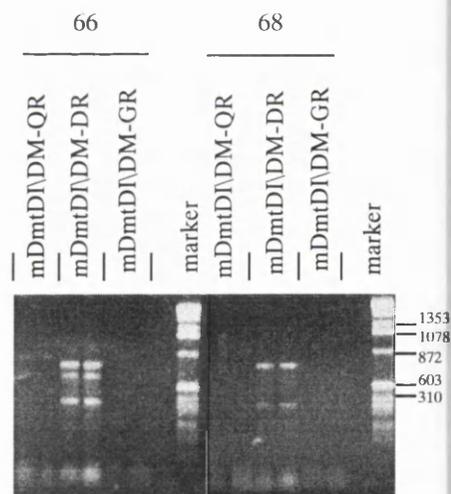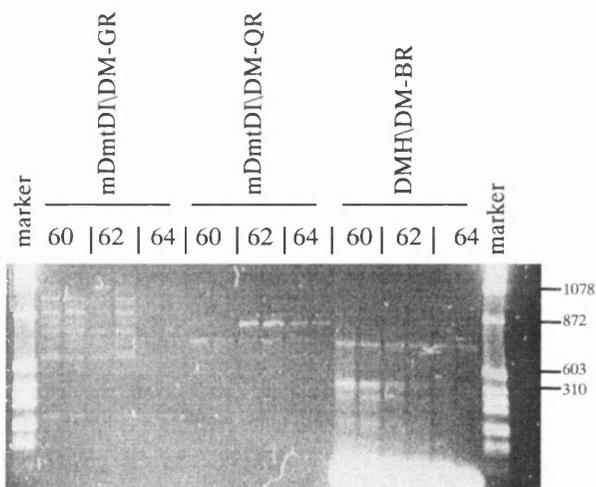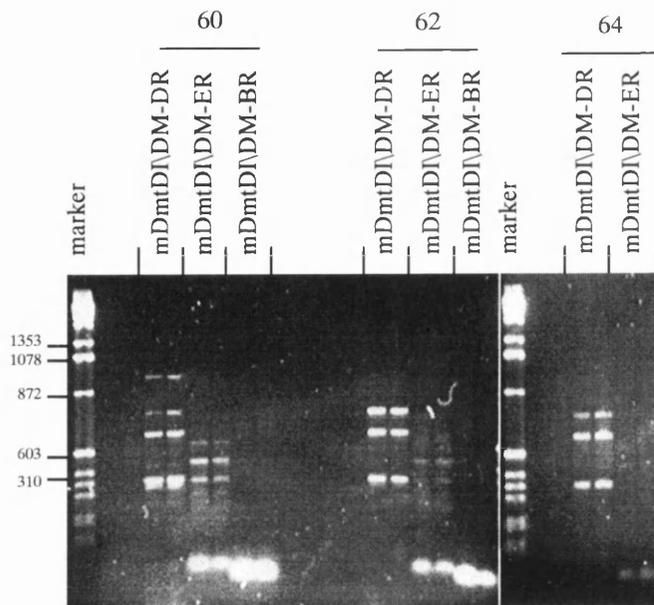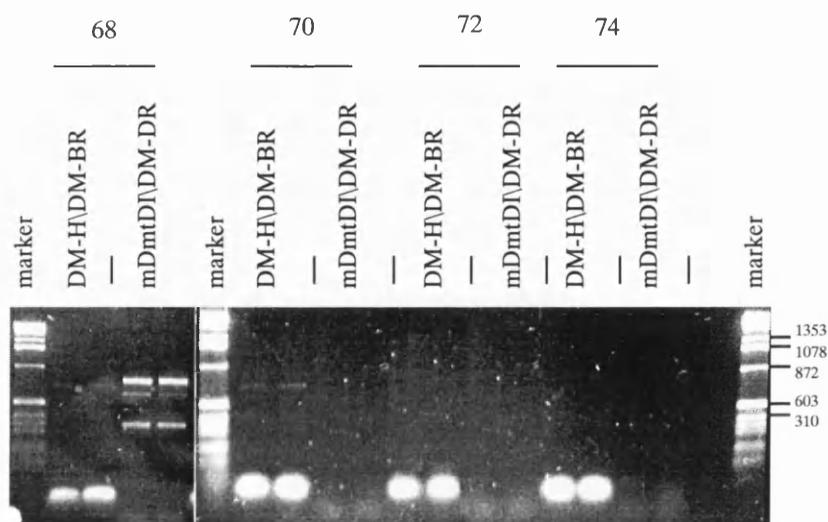
Figure 3.5
Combinations of oligonucleotide primers at a selection of temperatures
Each PCR amplification was carried out in a 7 µl volume that contained, ~ 1ng of DNA, 1 X buffer, the forward and reverse primers under test at a 1µM concentration and 0.175 units of *Taq*. The PCR reactions were cycled 30 times under standard conditions using the stated temperature. The amplified template DNA had a repeat size of 162. The DM-H / DM-BR oligonucleotide primer combination produces an 829 bp product, DM-H / DM-DR a 674 bp product, Dmt-DI / DM-BR a 749 bp product, Dmt-DI / DM-DR a 754 bp product, Dmt-DI / DM-ER a 700 bp product, Dmt-DI / DM-GR an 852 bp product and Dmt-DI / DM-QR an 1187 bp product. Each PCR product was resolved on a 1.5% agarose gel (w/v) and photographed.

that the same pair could be used to study all 4 lines.

It was determined that the PCR product of an amplification with DM-H / DM-DR would be around 670 bp in size. As the majority of the SP-PCR products resolved on the agarose gels were not visible until after hybridisation, the 603 bp band of the φ X174 *Hae* III digested marker was used as a guide to the location of the SP-PCR products. 10 cm of the gel above and below this band was transferred on to a nylon membrane in the first instance. This would encompass the gel between the 271 band and the 1353 band. Again assuming the PCR product to be 670 bp in size these boundaries would allow deletions of up to 133 repeats and expansions beyond 230 repeats to be visualised. If a particular tissue appeared to be unstable beyond this 10cm limit then this limit was modified appropriately.

3.3.3 Choice and preparation of tissues

Tissues were chosen for a selection of reasons. This included tissues known to be affected in DM1 patients, including the eye and brain as well as a selection of muscles. The nature of the cells that made up a tissue type was also taken into account. Autoradiography has been used to examine the rate of cell proliferation and cell migration (Messier and Leblond, 1960) by examining the uptake and migration of radiolabelled precursors. The pattern of the incorporation of the radiolabel into the nuclei and its subsequent dispersal was collated and used to classify the cell populations. They were categorised into three groups, these include:

Static; no radiolabelled nuclei were observed in such a population, indicating no change in the cell population.

Expanding; a small number of radiolabelled nuclei appear. These appear to be incorporated permanently into the cell population. This group includes all muscle cells and most connective tissues.

Renewing; a large number of radiolabelled nuclei appear and subsequently decrease, indicating cell loss. Such cells are being constantly replaced. These include cells found in the blood.

The tissues that were investigated and the reasons for that choice are summarised in table 3.1. Small pool PCR products were consistently amplified for all of the tissues under examination with the exception of blood. Although the preparation method included steps to remove the haem group from the pelleted leukocytes before they were lysed, it is possible that this was not always the case and the presence of haem within the leukocyte DNA could be inhibiting the

tissue	reason for choice	cell population (Messier and Leblond, 1960)
blood, spleen	Source of leukocytes; leukocytes are the most easily assessable tissue obtained from all patients	Leukocytes ; renewing.
Quadriceps femoris (skeletal muscle)*, heart*, colon*, diaphragm, tongue	DM1 is the only muscular dystrophy that affects non skeletal muscles	All muscle cells; expanding. Tongue / colon epithelium; renewing.
eye*	A common initial symptom in DM1 pedigrees is the development of cataracts	Retinal cells; static. Lens epithelial; expanding. Corneal epithelial; renewing.
brain*	Congenital DM1 patients exhibit mental retardation; many CAG•CTG tract disorders exhibit neurodegeneration	Neurones; static. Glial; expanding.
kidney	Observed to be unstable in other triplet repeat models	Glomoreli, proximal and distal tubules; renewing.
liver	Observed to be unstable in other triplet repeat models	Parenchymal and littoral cells; expanding
lung	An example of a high turnover tissue	All cells renewing
tail	Weaned tail used to define progenitor repeat size	

Table 3.1 Summary of none tail tissues examined in investigation of somatic instability. *Tissues effected in DM1 patients.

PCR amplification. However, the when the instability pattern of blood was present it was found to echo that of the spleen which also has a large leukocyte population. The spleen was consistently amplified and it was assumed that this tissue exhibited the behaviour of the repeat in the leukocyte. The tissues examined included tail DNA taken from each mouse shortly after weaning at ~ 6 weeks. The transgenic repeat size of weaned tail was taken to be the progenitor repeat size. A sample of this progenitor allele DNA was resolved on the same gel as the tissues taken after sacrifice again making subtle changes in repeat length detectable. Each tissue sample that was studied using small pool PCR was amplified at least 2 different concentrations, at approximately 1 molecule per reaction and around 100 molecules per reactions. The repeat profile of the lower concentration reaction always fell within the profile of the higher concentration reaction for the same sample.

3.4 Somatic mosaicism in *Dmt* mice

3.4.1 Somatic mosaicism in mature *Dmt* 162 transgenic mice

It is known that the unstable repeat of the disease allele expands over the lifetime of a DM1 patient (Wong *et al.*, 1995). The effect that the age of the mouse has on the transgenic repeat in the different tissues was initially investigated in 20 month old mice in *Dmt*-B, -C -D and -E (Figure 3.6). A male and female mouse was examined for each line to examine whether the sex of the mouse had an effect on the somatic instability observed. *Dmt*-B, -C and -E exhibited minimal instability, that is they did not show a length change beyond a +/- 20 limit in either sex. *Dmt*-D however, exhibited a significant expansion biased instability that varied according to the tissue in both sexes. Most notably the repeat was strikingly unstable in kidney and this instability was clearly expansion biased. Indeed a group of kidney cells contained as many as 650 repeats, which means that these alleles have gained in excess of 500 repeats from the progenitor repeat size. Other tissues that were particularly unstable included the liver, colon, skeletal muscle, diaphragm and eye. The repeat in the tail, spleen, heart and lung however, was relatively stable. There was no correlation observed between the level of stability and the level of turnover of cells populating that tissue. A high level of instability was observed in brain and eye which contain non dividing cells, while the constantly renewing cells in both the lung and the spleen demonstrated some of the lowest levels of instability observed.

3.4.2 Age dependence of somatic mosaicism in *Dmt*-D

To investigate whether the somatic instability observed in the tissues of the mature *Dmt*-D mice progressed over the lifetime of the mouse, DNA was extracted and examined from mice
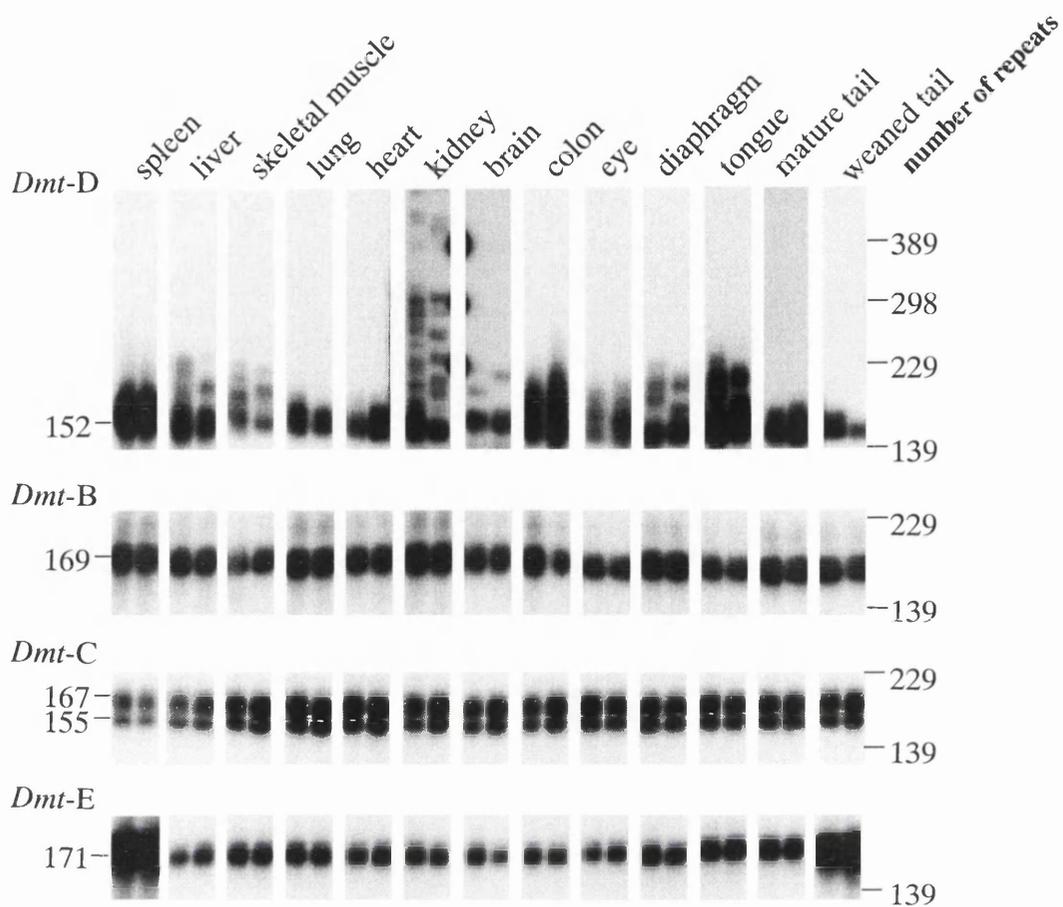
Figure 3.6
Somatic variation in mature *Dmt* transgenic mice
Multiple tissues were taken from 20 month old male mice, these included a *Dmt*-B with a transgenic repeat size of 169, a *Dmt*-C with transgenic repeat sizes of 155 and 167, a *Dmt*-D with a transgenic repeat size of 152 and a *Dmt*-E with transgenic repeat size of 155. Between ~10 and 100 molecules from each tissue were used as the template for duplicate small pool PCR amplifications which were resolved on a 1.25% (w/v) agarose gel before transfer to a nylon membrane and Southern hybridisation with DM56. The molecular weight marker has been converted to indicate the number of triplet repeats.

at 2, 6 and 13 months (Figure 3.7). At 2 months of age all of the examined tissues appear to be relatively stable with the exception of the kidney and colon, tissues which had been observed to be particularly unstable in the mature mouse. At this stage some large expansion events, beyond 20 repeats have already occurred in the kidney. The tissues at 6 months reveal the instability to be progressing over time, and this progression was confirmed when the tissues were studied at 13 months. However, the rate of this expansion
appeared to vary according to the tissue tested. As the *Dmt*-D mice were investigated there appeared to be a tissue specific pattern of repeat instability. To confirm this the tissue selection was compared in three male and three female mice 12-13 months of age. The instability pattern was consistent from mouse to mouse, with the kidney persisting as the tissue demonstrating the largest variation in repeat length. Comparison of up to 10 molecules of kidney DNA for each small pool PCR reaction at each timepoint (Figure 3.8) reveals an age dependent pattern of mosaicism.

3.4.3 Single molecule analysis of somatic mosaicism in the lung and kidney of *Dmt*-D mice

To allow quantification and analysis of the pattern of stability of the unstable kidney and relatively stable lung of *Dmt*-D mice at each of the different timepoints, multiple SP-PCR amplifications (10 or 11) were carried out with each concentration as the template, either 1200 $pg\mu l^{-1}$, 120 $pg\mu l^{-1}$ or 12 $pg\mu l^{-1}$. These preliminary reactions were used to determine more accurately the number of molecules in each lane produced at the lower concentrations using the Poisson distribution. As the concentrations have been derived by dilution, lanes which contain molecules may contain more than one molecule of similar size thus using positive lanes to calculate the number of molecules may under estimate the concentration. However, the blank lanes, (p(0)), all lack one molecule therefore p(0) is calculated as:

total number of lanes - number of positive lanes / total

Calculation of p(0) allows the number of molecules present in each SP-PCR amplification at this concentration to be derived from the Poisson distribution where

number of molecules per reaction = -ln p(0)

71
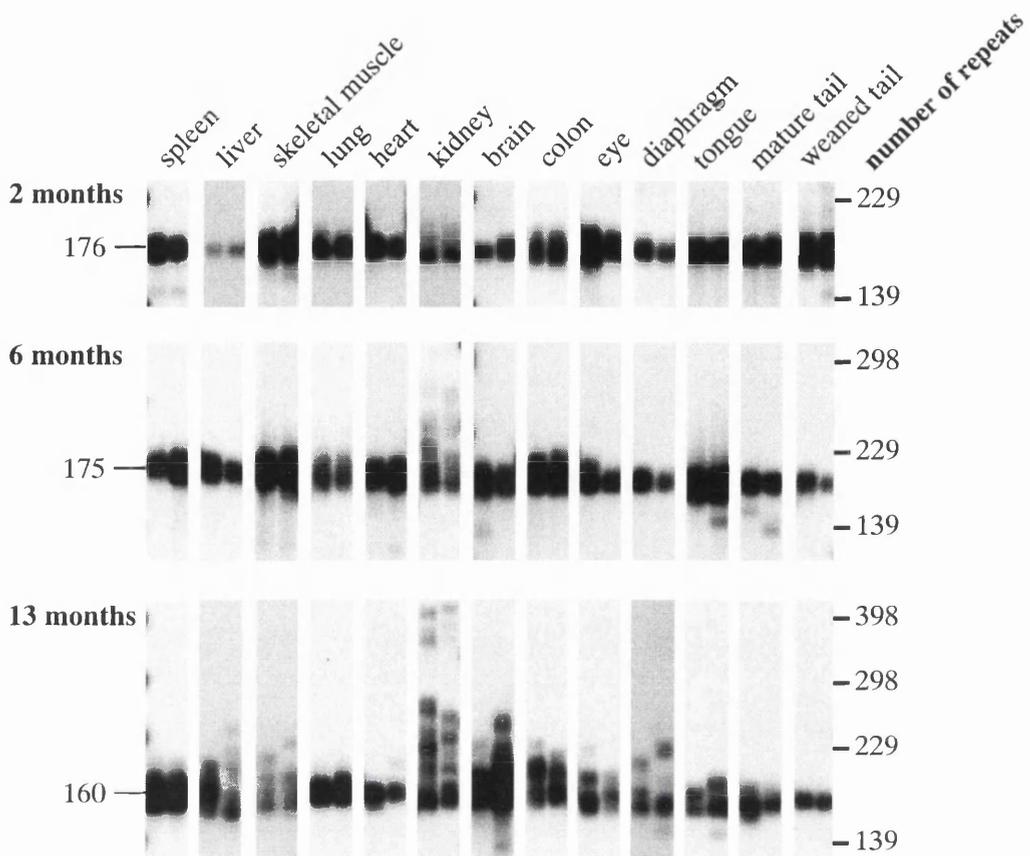
Figure 3.7
Somatic variation in *Dmt*-D transgenic mice increases with age
Multiple tissues were taken from *Dmt*-D male mice at 2 months, 6 months and 13 months, with a progenitor allele size of 176 repeats, 175 repeats and 160 repeats respectively. Between ~10 and 100 molecules from each tissue were used as the template for duplicate small pool PCR amplifications which were resolved on a 1.25% (w/v) agarose gel before transfer to a nylon membrane and Southern hybridisation with DM56. The molecular weight marker has been converted to indicate the number of triplet repeats.

From these calculations it was possible to carry out reactions on single molecule templates. Around 100 molecules was taken as the minimum number that should be examined for each mouse at each time point. Also at least 32 SP-PCR amplifications were carried out with 100 molecules of DNA from each tissue and time point as the template. As well as giving an indication of the nature of the distribution of the repeat length (Figure 3.4.D), it also meant that rarer expansion events could also be observed. For each tissue the same stock of master mix was used, with the appropriate amount of DNA added to achieve the correct number of molecules for that reaction a set of negative reactions using water in lieu of DNA was also set up with the stock mastermix to detect external contamination. Study of the pattern of instability in lung cells at each time point found a steady rate of expansion biased events, and that the average size of the alleles steadily increased over time. At 2 months when compared to the progenitor allele length (weaned tail DNA) it was found that more than 30% of the kidney cells analysed had increased in allele length by more than 10 repeats. However, the same distribution revealed that ~ 7% of the cells had deletions greater than 5 repeats. These data indicate that although there is an overall bias toward expansion of the transgenic repeat, a portion of the expansion events result in loss of repeats. By 13 months the kidney distribution contains repeats that are more than 200 repeats larger than the progenitor allele. The expansion pattern within the kidney does not appear homogenous, in fact the repeat distribution suggests that there are at least three different modes of repeat expansion within the kidney. The most stable group of cells, which make up ~35% of the molecules studied, exhibit expansions of ~ 7 repeats from that of the progenitor allele (peak I). The largest proportion of the cells, ~ 40% exhibit mean gains of ~ 50 repeats from the progenitor allele (peak II), whereas the ~25 % of cells which had gained the largest increase had expanded ~150 repeats beyond the progenitor size (peak III). These modes, which first became apparent at 13 months are even more striking at 20 months, with peaks I, II and III exhibiting increases of around +10, +100 and +270 respectively. When approximately 200 molecules of kidney DNA at 13 and 20 months were used as the template for each SP-PCR amplification (Figure 3.9), the trimodal distribution of the cell population can be seen clearly on the resulting Southern hybridisation (Figure 3.10). In contrast, this multimodality was not observed in the lung, although there was a gradual increase in the allele length at each timepoint.

3.4.4 Somatic mosaicism in a 25 month old *Dmt*-E mouse with a tumour

A 25 month old *Dmt*-E female was sacrificed because it had developed multiple muscle sarcomas. This and subsequent tumours were preliminarily identified and classified by Chuing-Mei Cheng, MD, University of Glasgow. DNA was prepared from a selection of
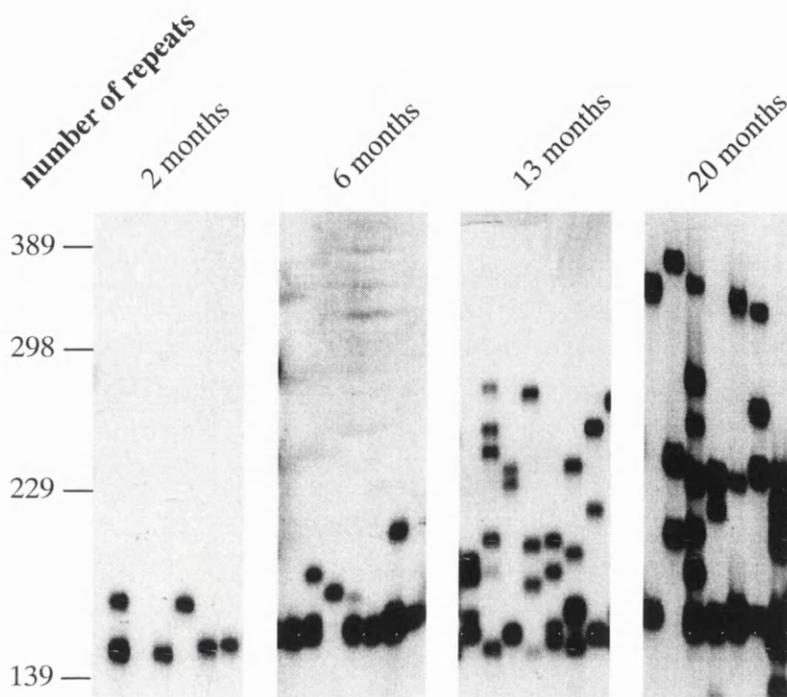
Figure 3.8
SP-PCR analysis of age dependent variation in kidney fom *Dm*t-D mice
Kidney DNA from *Dmt*-D male mice aged 2, 6, 13 and 20 months with a transgenic repeat size of 176, 175, 160 and 156 repeats respectively was used as the template for small pool PCR amplifications. Between 1 and 10 molecules of DNA were used as template and a representative selection of reactions are shown. The molecular weight marker has been converted to indicate thenumber of triplet repeat
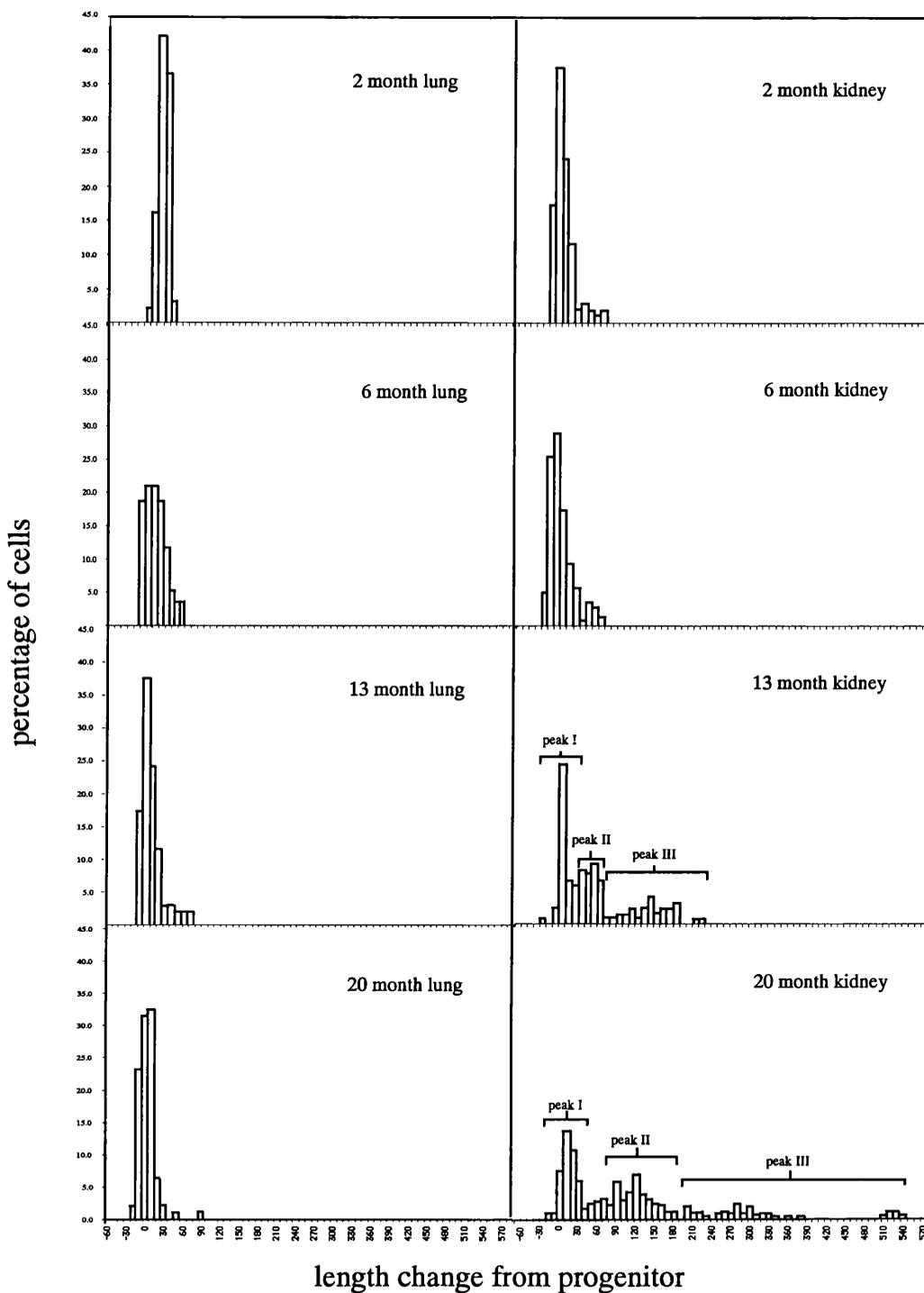
Figure 3.9
Quantification of age-dependant variation in lung and kidney from _Dmt-D_ mice
Histograms of repeat length distibutions relative to the progenitor allele in the lung
and kidney of male _Dmt_-D mice aged 2, 6, 13 and 20 months with a transgenic repeat
size of 176, 175, 160 and 156 repeats respectively was used as the template for small
pool PCR at low DNA concentrations (1-5 molecules). Allele lengths have been grouped
into 10 repeat size ranges. The three peaks of variability within the kidney cell
distributions derived from the older mice are also shown.

75

tissues, including the muscles the tumours appeared to have developed from as well as the tumours themselves. The DNA was diluted to 1200 pgμl^{-1}, 120 pgμl^{-1}, 12 pgμl^{-1} in dilution buffer containing 0.1μM DM-H. The 1200 pgμl^{-1} and 12 pgμl^{-1} concentrations were used as templates for duplicate SP-PCR amplifications. Tissue by tissue comparison revealed incidences of deletion biased instability in the tumour derived DNA greater than anything previously observed in *Dmt*-E tissues (Figure 3.11). Examination of the repeat in the heart, lung, colon, kidney tail and skin found it to be stable, whilst instability was observed in DNA derived from the blood and part of the muscle which one of the tumours had been associated with.

3.4.5 Somatic mosaicism in *Dmt* 162 transgenic mice using Genescan software®

Genescan® software has been used to visualise the instability profile of different tissues in several mouse models of expanded triplet repeat instability (Mangiarini *et al.*, 1997; Sato *et al.*, 1999; Wheeler *et al.*, 1999). To examine further the instability profile of *Dmt*-D mice, 1 μl of tail DNA lysate (~1ng) was used as the template of a PCR amplification which contained 1 X PCR buffer, 0.25 Units of *Taq*, 5% DMSO and DM-H and DM-DRF (a fluorescently labelled Genescan® primer) in a final concentration of 1μM in a 10 μl volume. Each reaction was cycled 30 times under standard conditions, with an annealing temperature of 68 °C and resolved on a 6% acrylamide gel (v/v). The conditions and precautions taken while preparing samples for Genescan® analysis are discussed in Chapter 5.

Around 1ng of DNA was amplified for a selection of *Dmt*-D tissues taken from the 20 month old male *Dmt*-D mouse in (3.4.1). Instability profiles distinct to each tissue type were observed and the pattern of instability levels of the different tissues relative to each other were in common to those observed on SP-PCR (Figure 3.6). However, although the kidney tissue is found to be the most unstable of all of the tissues, the level of instability observed is not as great as that detected in the same tissue using SP-PCR and indeed can be detected at much lower concentrations. Increases greater than 500 repeats from the progenitor were observed in the 20 month old *Dmt*-D kidney using SP-PCR. Repeats of not larger than 80 repeats were detected Genescan® software (Figure 3.12) to detect repeat expansions in a higher concentration of template DNA of the same sample. This lack of sensitivity to the level of instability present in a sample was a consistent observation when using Genescan® analysis.

76
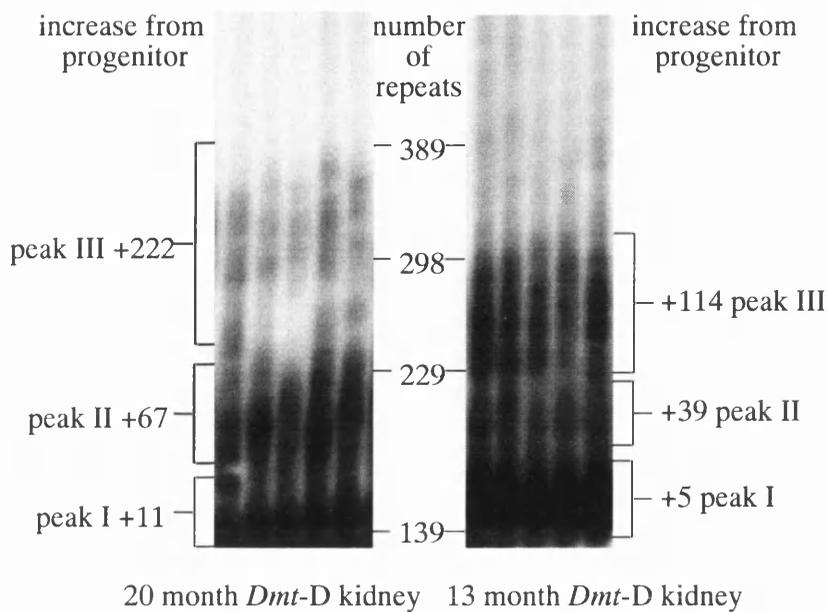
Figure 3.10
Multimodality of repeat length disribution in kidney cells from *Dmt*-D mice
Approximately 200 molecules of kidney DNA were used as the template for small pool PCR amplification to demonstrate the multimodality present in the kidney cell population of mature male *Dmt*-D mice. The 20 month old mouse has a transgenic repeat size of 156 and the 13 month old mouse has a transgenic repeat size of 160. The molecular weight marker has been converted to the number of triplet repeats.
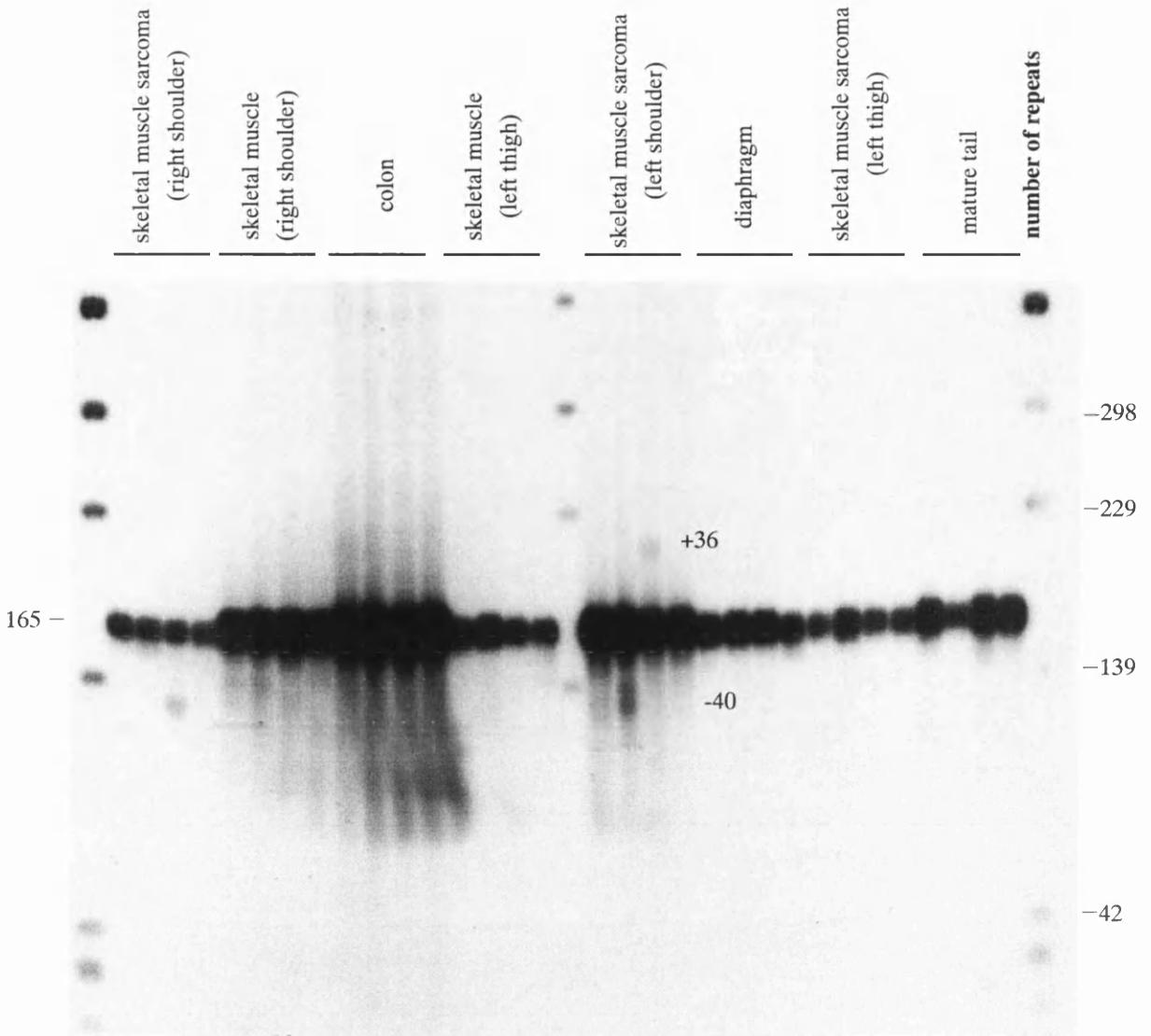
Figure 3.11
Somatic mosaicism in a 25 month old *Dmt-E* mouse with multiple tumours
DNA from a 25 month old *Dmt*-E female with a repeat size of 165 was extracted from multiple tissues including muscle sarcomas and the muscles associated with the tumours. Between 50 and 100 molecules of DNA were used as the template for quadruplicate SP-PCR amplifications, the PCR products were resolved on a 1.25% agarose gel (w/v) before transfer to a nylon membrane and Southern hybridisation with DM56. The marker has been converted to show the number of repeats. Expansions / deletions have also been indicated.

78
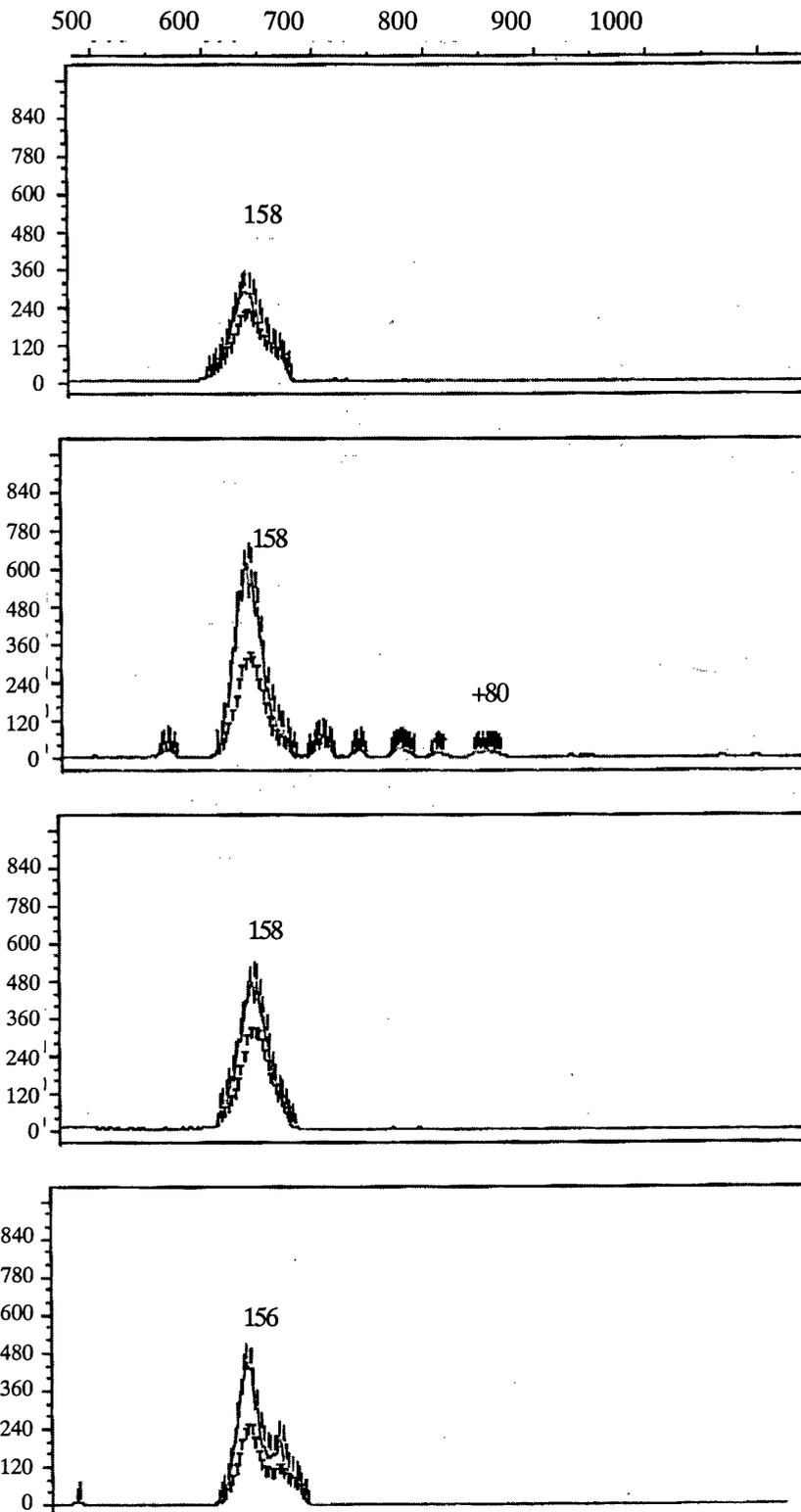
Figure 3.12
Genescan profiles of tissues from a 20 month old male mouse
Approximately 1ng of the A; weaned tail, B; kidney, C; mature tail, D; colon,of a 20 month old *Dmt*-D mouse with a 156 repeat progenitor size was used as the template in a PCR amplification which contained 1 X PCR buffer, 0.25 Units of *Taq*, 5% DMSO and DM-H and DM-DRF at a final concentration of 1mM in a 10 µl volume. 1.5 µl of a 1:1 mixture of PCR product and Genescan® mix was denatured at 90°C before resolving products on a 6% acrylamide gel.

79

3.5 Regional mosaicism in *Dmt*-D mice

3.5.1 Regional mosaicism in the kidneys of *Dmt*-D mice

The instability events observed in *Dmt*-D appear to be the accumulation of multiple small events, the majority of which are expansion biased. It has become apparent that the transgenic repeat within cells that make up the kidney population are subject to a different amount of repeat expansion events. That is different cells have a different rate of expandability. Are these different cell populations located in different regions within the kidney?

To investigate and possibly identify what these cells might be, kidneys from mature *Dmt*-D mice, 12, 15 and 30 months were dissected out and placed into a petri dish containing ~50 µl 1 X SSC. They were placed under a dissecting microscope or magnifying glass. The outer membrane, the capsule, was removed before they were sliced transversely and succinct punches removed from distinct regions (Figure 3.13.A). Discrete punches from each region were removed and placed into 50 µl of proteinase K lysis buffer in a 200 µl PCR tube and incubated overnight at 60 °C. The tissue was diluted 1/100 in distilled water and the proteinase K heat inactivated.

DNA generated from these regions was used as a template for SP-PCR analysis and reactions were carried out in quadruplicate. The samples were resolved along side total kidney and weaned tail DNA (Figure 3.13.B).

It appeared that the alleles with the most unstable repeat length were located in the cortex of the kidney. Investigation of cell types within this region found it to contain highly specialised structures such as the glomereli and tubules, which include cell types that were not present in the rest of the kidney. An attempt was made to isolate specific structures within the kidney.

The kidneys were removed from mice aged 15 and 30 months and placed in a petri dish containing 1XPBS before slicing transversely and placing in an Eppendorf that contained enough tissue culture buffer to cover the kidney. The kidney was then snap frozen in liquid nitrogen until needed, when the tissue was allowed to defrost to an optimal temperature, ~ -18 °C.

The kidney was sliced at 30 microns, the finest setting available on the AS 620 Cryotome, (Anglia Scientific Instruments Ltd.) that was used. Every fifth slice was placed onto gelatine coated slide. The kidney slices were dissected under 1 X PBS using a microscope. Succinct structures, including tubules (the proximal and distal tubules could not be separated) and glomereli were isolated and placed in 5 µl of PBS in 200 µl PCR tubes. The structures were removed using a combination of fine gauge needles and a mouth pipette. The cells were lysed

A

capsule
cortex
glomerulus
nephron
pelvis
calyx

B

number of repeats

881—
171—

cortex
pelvis
calyx
capsule
total kidney
weaned tail

—389
—298
—229
—139
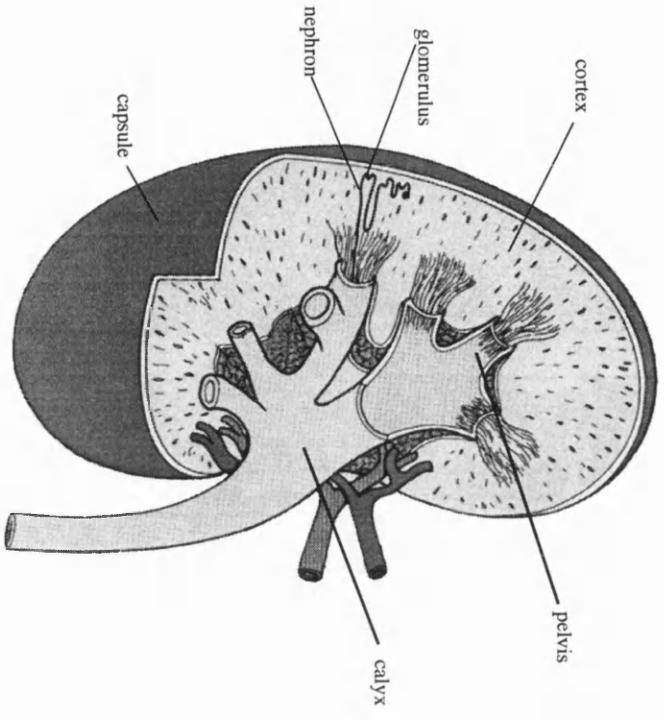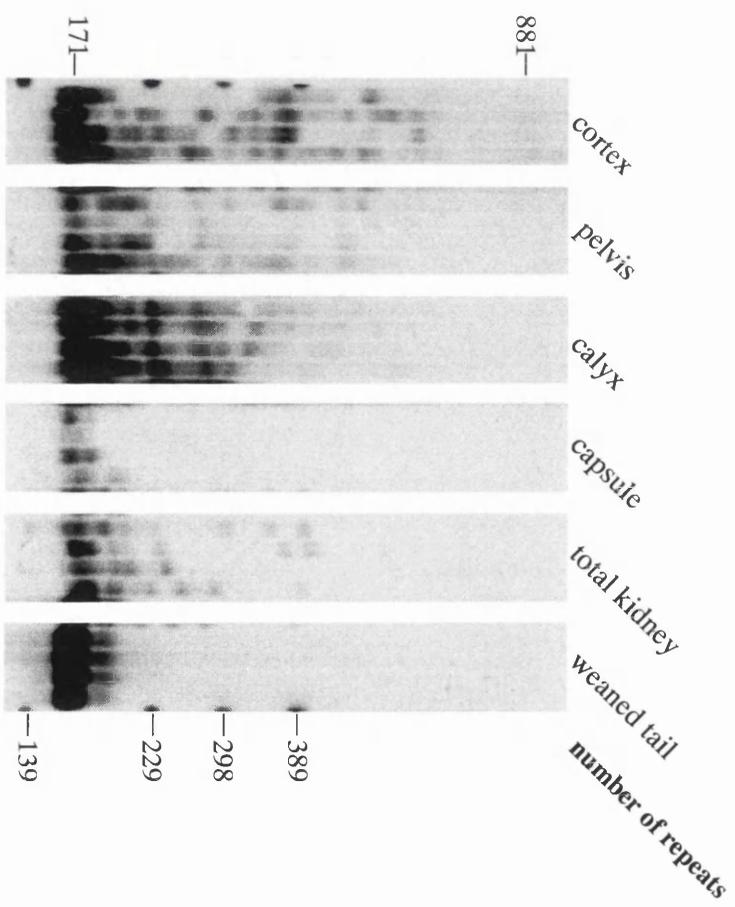
Figure 3.13
Regional mosaicism in the kidneys of *Dmt*-D mice
A. Diagram indicates the location of the regions of the kidney that were sampled for SP-PCR analysis.
B. DNA was prepared from the cortex, pelvis, calyx, capsule and total kidney of a 15 month old mouse with a progenitor repeat size of 171.
The DNA was used as the template for quadruplicate SP-PCR amplifications which were resolved on a 1.25% (w/v) agarose gel before transfer
to a nylon membrane and Southern hybridisation with DM56.

using 5µl KOH lysis buffer and incubated for 10 minutes at 60 °C before neutralising in 5.2µl 200mM tricene. All of the lysate was used as the template for the SP-PCR amplification of the repeat that was carried out in a 30µl volume under standard conditions for DM-H and DM-DR. In the first instance 1 µl of this PCR product was used as the template for a second round of SP-PCR amplification in a 7 µl volume, under standard conditions for DM-H and DM-DR. The PCR products from each round of amplification were resolved on a 1.25% (w/v) agarose gel before transfer to a nylon membrane by squash blotting (Figure 3.14). Not all of the isolated structures were resolved as PCR products. It was thought that the failure of some of the PCR reactions could be due to the failure of the isolated cells to lyse properly to release the DNA template. An attempt was made to increase the efficiency of the lysis of the discrete structures by lysing the cells in 10 µl of proteinase K lysis buffer for 10 minutes and briefly heat inactivating the proteinase K before using the total lysate as the template for SP-PCR amplification. This method of lysis appeared to be no more efficient than the KOH lysis. As this series of experiments progressed, the presence of the structures within the thin walled tubes was confirmed under the microscope.

The structures that were resolved did not allow conclusions to be drawn about the identity of the cell type or types that were responsible for most dramatic changes in allele length. Comparison of glomereli and tubules were inconclusive.

3.5.2 Regional mosaicism in the brains of *Dmt*-D mice

Given that the majority of CAG•CTG repeat expansion disorders are neurodegenerative diseases and that the brain in *Dmt*-D was one of the more unstable tissues an investigation of repeat mosaicism within the brain regions was carried out.

The brain of a 12 month old *Dmt* D mouse was completely dissected out of the skull and placed in 1 X PBS in a petri dish. Microdissection of the hippocampus, cerebellum, striatum, cortex and olifactory bulb were carried out under a dissecting magnifying glass before each region was placed into 500 µl of proteinase K lysis buffer and incubated overnight at 60 °C. The microdissection of the brain was carried out by an Honours student, Mr Christos Vassilopoulos under the guidance of Dr Heather Johnstone of the University of Glasgow. Serial dilutions of 1/10 (~ 1000 pgµl^{-1}), 1/100 (~ 100 pgµl^{-1}), 1/1000 (~ 10 pgµl^{-1}) were carried out in dilution buffer containing 0.1µM DM-H. 1µl of the 1/10 and 1/1000 serial dilutions were used as templates for duplicate small pool PCR amplifications for each region after the

Figure 3.14
Isolation of discrete structures of mature *Dmt*-D kidney
Structures were isolated from a 30 micron *Dmt*-D kidney slice taken from a 15 month old female with a repeat size of 161. The cells were lysed with 200mM KOH and neutralised with 200 mM tricene. All of the lysate was used as the template of a SP-PCR reaction in a 30 μl volume in the first round of PCR. 1 μl of of the first round of PCR product was used as the template in the second round of SP-PCR amplifcation in a 7 μl volume. The PCR products were resolved on a 1.25 % agarose gel (w/v) and transferred onto a nylon membrane before hybridisation with DM56.
A Autoradiograph after 3 hours exposure. B Autoradiograph after 16 hours exposure.

83

proteinase K was heat inactivated. The samples were resolved alongside weaned tail DNA (Figure 3.15.A). Clear differences in the degree of stability were observed in the different regions with a high level of variation observed in the hindbrain and the striatum while a low level of instability was observed in the cerebellum.

Quantification of the degree of stability of the cerebellum and striatum was carried out. DNA was prepared by phenol / phenol: chloroform extraction. The concentration of the DNA was determined and diluted to 1200 pgμl^{-1}, 120 pgμl^{-1} and 12 pgμl^{-1} and the repeat length of single molecules determined (Figure 3.15.B).

It was observed that the mean increase of repeat length in the striatum was +25 repeats and that although the cerebellum was relatively stable there was a shift of the average allele size from that of the progenitor repeat size of +10 repeats.

3.6 Conclusion

The aim of the work in this chapter was to establish the suitability of the *Dmt* transgenic lines as models for triplet repeat instability. The investigation of the 4 lines found that one, *Dmt*- D replicates the gross somatic mosaicism observed in DM1 patients, with some cells exhibiting expansions beyond 500 repeats which is 3 times the progenitor allele size. Although it is not possible to compare the tissue instability spectrum of *Dmt*-D to the same tissues in DM1 it would appear that the kidney of *Dmt*-D is as unstable as tissues observed in DM1 patients. This would suggest in the first instance that the mouse metabolism can display a mechanism of instability parallel to that observed in man and in the second that in the *Dmt*-D triplet repeat model, we have a model suitable for investigating mechanisms of triplet repeat expansion. Comparison of the level of cell turnover within the different tissue types to the pattern of stability found no correlation between the two. Indeed the level of cell turnover of the kidneys is much lower than that of the lung, one of the most stable tissues (Messier and Leblond, 1960). Also there is no evidence of repeat instability in the very young mice despite the rapid rate of cell division during embryogenesis, again a parallel can be drawn between DM1 as little somatic instability is observed in DM1 newborns (Tachi *et al.*, 1993). Other evidence against the rate of cell proliferation having an effect on the level of instability can be seen in the microdissection of the brain. Analysis of the regions of the brain exhibit a regional specific
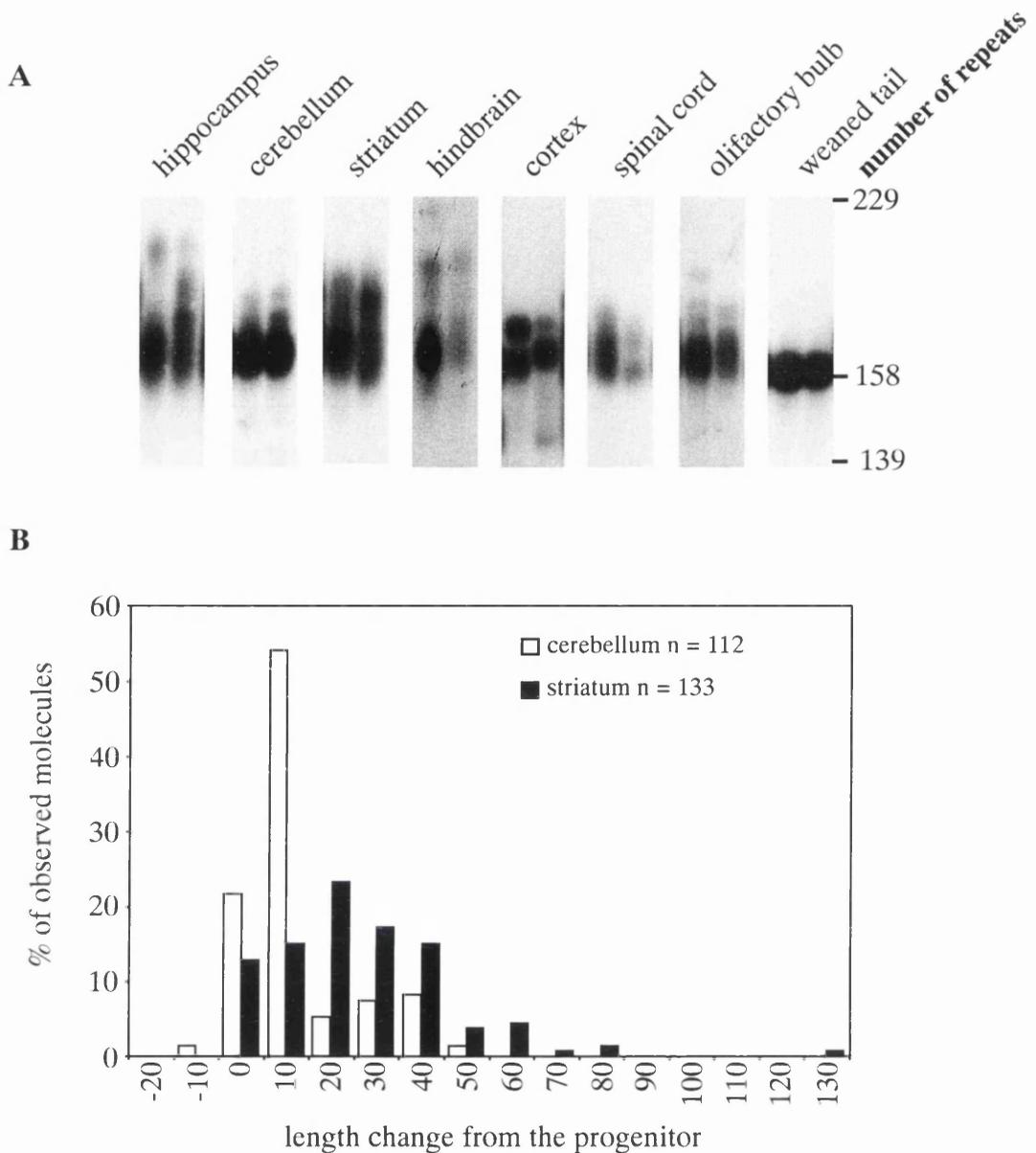
Figure 3.15
SP-PCR analysis of regional dependant somatic mosaicism in the brain of a *Dmt*-D mouse
A Small pool PCR reactions were carried on out on regions of the brain of a 12 month old *Dmt*-D male mouse with a repeat length of 158. Between ~10 and 100 molecules from each tissue were used as the template for duplicate amplifications. The products were resolved on a 1.25% agarose gel (w/v) and transfered onto a nylon membrane before Southern hybridisation with DM56. The molecular weight marker has been converted to indicate the number of triplet repeats.
B Histograms of repeat length distribution of the cerebellum and striatum of the male mouse shown in A. The molecule length was derived by SP-PCR amplification at low concentrations (1-5 molecules).

85

pattern of instability including increases beyond 80 repeats from the progenitor observed in the striatum. Yet the neuronal cells are post mitotic, indeed although the majority of brain cells are glial, the rate of division of these cells is at a rate of less than once in a mouse older than four months (Cameron, 1970; McCarthy and Leblond, 1988). This data suggests that the repeat expansions are accumulating in nondividing cells.

Single molecule analysis of the kidney has revealed multi-modality in the allele distribution with at least 3 distinct cell populations displaying different repeat dynamics. Examination of the instability in the different regions again revealed a regional specific pattern of mosaicism. This and indeed the brain data suggest, perhaps not surprisingly that the repeat instability is not only dependent on the tissue type but in fact the type of cell. The kidney cell population includes nervous, lymph and vascular system in common with tissues exhibiting more stable repeat dynamics suggesting that the dramatically unstable cell population does not have its origin here but within a highly specialised cell in perhaps the nephron. Identification of these cell types may allow further insight into the mechanism of instability yet methods to investigate such cellular differences *in vivo* will require further development.

That the other lines, *Dmt*-B, -C and -E exhibited minimal instability highlight the role played by the integration site on the instability levels of the transgenic repeat tracts. It is known that transgenes may be subject to silencing effects at some integration sites (Clark, 1997), or have integrated into regions which were already rich in chromatin. The transgene itself may have been subject to mutation, a point mutation being enough to reduce the purity of the tract. The SCA type 1 CAG repeat in normal range individuals contains a CAT repeat which is not present in expanded alleles and is thought to confer a stabilising effect (Jodice *et al.*, 1994).

However, the instability observed in the otherwise stable line *Dmt*-E after it had developed tumours suggests that other factors also play a role repeat instability. Recalling that each individual within a *Dmt* line is identical other than the length of the repeat tract, factors effecting the stability of the repeat in one of the stable lines have been acquired by the individual over their lifetime. The deletions observed in the *Dmt*-E mice occurred either within the tumours themselves or the tissue from which the tissue had been derived. This suggests a mutation event has occurred which effects the integrity of the repeat.

Chapter 4

Gametic instability in the *Dmt* transgenic lines

4.1 Introduction

The analysis of DM1 pedigrees has revealed that the size of the repeat transmitted from a DM1 parent to its offspring is dependent on the repeat size carried by the parent as well as the parents sex (Ashizawa *et al.*, 1994; Brunner *et al.*, 1993; Harley *et al.*, 1993; Jansen *et al.*, 1994; Lavedan *et al.*, 1993a). In pedigree studies based on leukocyte DNA, more than 90% of transmissions resulting in an increase in the repeat length observed in the offspring and only 6-7% resulting in a decrease (Ashizawa *et al.*, 1994; Brunner *et al.*, 1993). Monckton *et al* (1995), Martorell *et al* (2000) carried out intergenerational studies and examined the pattern of instability in germline DNA as well as the leukocyte DNA of DM1 males. The somatic cells studied revealed a bias towards increasing allele length and a lower boundary to the variation shown by the disease range repeat. Events producing alleles smaller than this boundary were rare. Examination of sperm DNA revealed extensive expansion variation in the disease range allele, with a pattern distinct from that observed in the soma but consistent with the phenomenon of anticipation observed in DM1 transmissions. There was little correlation between the progenitor repeat length and the variation observed in the sperm. Within the pool of germline alleles repeat contraction events were also observed including reductions of the disease range allele into a normal repeat range but such dramatic events were rare. The majority of observed deletions from the progenitor repeat size are likely not to be deletions. It is more likely that when determining the progenitor repeat size in an individual, repeat events have already started to occur in the tissue shifting the mean, and that the mutation rate in the tissue which the progenitor rate has been determined from (leukocyte) is higher than in the germline.

The sex of the transmitting parent influences the size of the repeat transmitted. It has been observed (Bell, 1948), and confirmed with the identification of the associated mutation (Aslanidis *et al.*, 1992; Brook *et al.*, 1992; Buxton *et al.*, 1992; Fu *et al.*, 1992; Harley *et al.*, 1992; Mahadevan *et al.*, 1992), that the transmitting grandparent in DM1 pedigrees is most often male. Indeed when the transmission event results in a small increase of not more than 100 repeats, the transmitting parent is more likely to be male (Brunner *et al.*, 1993; Lopez de Munain *et al.*, 1995; Wieringa, 1994). While maternal transmissions account for the larger expansions (Harley *et al.*, 1993; Lavedan *et al.*, 1993a; Redman *et al.*, 1993) and account for

the majority of transmissions which result in cDM offspring (Redman *et al.*, 1993; Tsilfidis *et al.*, 1992). This suggests a gametic mechanism of expansion dependent on differing sex dependent factors. However, it is possible that in addition to or perhaps instead of, expansions occur in a newly formed zygote and that these expansion events may be influenced not exclusively by the sex of the transmitting parent but by the sex of the gamete (Kovton *et al.*, 2000). Studies by Kovton *et al* (2000) on male transmissions of R6/1 Huntington's model mice has found that the male offspring are more likely to carry repeats larger than their parent while females were more likely to carry deletions.

Examination of the *Dmt* repeat transmission patterns in young male and female mice, found gametic instability present in all five of the *Dmt* lines, *Dmt*-A, -B -C, -D and -E and that the instability pattern exhibited varied depending on the sex of the transmitting parent repeats (Monckton *et al.*, 1997). That is maternal transmissions which did result in a repeat size different to that of the transmitting parent were deletion biased and could result in offspring with a repeat size as many as 7 repeats smaller than that of the parent. Germline mutation rates across the females of all of the *Dmt* lines was consistent at around 60%, and with exception of the multi integrant transgenic line *Dmt*-A, the mutation events were made up of small deletion events. Paternal transmissions were expansion biased and could result in offspring, which carried a repeat different to that of the transmitting parent by as many as +7. The mutation rate in male transmissions was found to be line dependent, from less than 10% in *Dmt*-E to almost 70% in *Dmt*-D.

The transmission pattern of the *Dmt* heterozygotes was examined for segregation distortion. That is the *Dmt* transgene was present in significantly more births than one in two. Segregation distortion has been observed in favour of the larger CTG repeat in DM1 at a rate of 56.4% (Carey *et al.*, 1994) and 58.1% (Hurst *et al.*, 1995). Analysis of the rate of the transmission of the *Dmt* repeat in the different lines, found the transgene present in 52 % of the births in *Dmt*-A, 50% of the births in *Dmt*-D and 46% of the births in *Dmt*-E these rates of transmission did not deviate significantly from the norm. However, significant segregation distortion in favour of the transgene did appear to be taking place in *Dmt*-B and *Dmt*-C with transgene transmission rates of 66% and 70% respectively (Monckton *et al.*, 1997). These transmission rates were for all births and not separated according to the sex of the parent.

The effect of age of the parent on the size of the repeat transmitted has been observed in paternal transmissions of Huntington's disease and DRPLA (Duyao *et al.*, 1993; Sato *et al.*, 1999). The maternal transmissions of a targeted *Sca* type 1 model also exhibited an age of parent at transmission effect, with the average size of deletion transmitted to the offspring

increasing as the mouse ages (Lorenzetti *et al.*, 2000).

An expanded CAG repeat originating from a SCA type 1 patient was used to generate two mouse lines, one D02, contained (CAG)82, the other, C01, (CAG)30 (Kaytor, 1997). Intergenerational instability in these lines was observed only in maternal transmissions, and these were deletion biased. Transmission studies observed that the transmitted deletion increased with the age of the mother. Furthermore examination of unfertilised oocytes at different timepoints found that oocytes isolated from a seven week old female to carry the same repeat size as the female. However, oocytes isolated from a 20 week old transgenic female found the oocytes to carry a repeat size smaller than that of the transmitting female.

4.2 Transmission studies

For the lines *Dmt*-D and *Dmt*-E, the size of the transgenic repeat carried by the offspring relative to that of the transmitting parent was determined over the lifetime of both transgenic male and female parents. All of the mice were on a FVB/n background and each mating was set up between a parent hemizygous for the transgene *Dmt* (+/-) while the other was wild type *Dmt* (-/-). This meant that all of the subsequent mice, if transgenic, were hemizygous for the transgene. The transgenic mice were identified as in Chapter 3.2 and the repeat size difference between parent and offspring determined by resolution of the fluorescently labelled transgenic repeat PCR product on a 6% acrylamide gel.

FVB/n was chosen as the strain on which the *Dmt* lines are based was due to several factors, it has a fully inbred genetic background and was reported to have a vigorous reproductive performance resulting in large litters. The fertilised eggs have large and noticeable pronuclei which favour the microinjection of DNA (Festing, 1998) and which have a higher level of survival than C57/BL6 (Taketo *et al.*, 1991). We found the FVB/n mice to be aggressive. This aggression was observed mostly in males although it was also seen in females. Littermates were often found to develop aggressive behaviour shortly after weaning. At its most extreme the resulting fighting would require the sacrifice of the whole cage of mice. There was no line specific pattern to the aggression observed. Several measures were taken to alleviate the aggression. These included the placing of the mice into fresh and not just clean cages each time the cage was due to be cleaned. This was found to have little effect. The use of larger cages and "toys" to alleviate boredom were judged to increase aggression. On a more basic level, the aggressive mice were housed singly. However, the most affective measure taken was the splitting of the colony into 2 separate rooms, the first housed only males, the second

housed the females and breeding cages. This reduced the level of fighting observed. Another unusual behaviour was the constant, unidirectional circling of certain mice. This behaviour was neither line nor sex specific. A consequence of these behaviours was that a large proportion of breeding pairs set up failed to result in live young. More than half of the transgenic females in breeding pairs failed to produce live young. The rate of successful paternal transmissions of the repeat was higher than that for maternal transmissions because a transgenic male can be paired with multiple wild type females, and only one of these combinations was required to be successful. As the success of specific matings was unpredictable attempts to focus on matings involving particular repeats i.e. larger or smaller than average were unsuccessful.

4.3 *Dmt* transgene allele sizing by fluorescent PCR using Genescan® software

The number of *Dmt* transgenic repeats carried by each transgenic mouse was determined using Genescan® software on an Applied Biosystems 373 A sequencer by the University of Glasgow's Molecular Biology Support Unit. The repeat size of each transgenic mouse was determined on a gel that also contained the repeat of the transmitting parent. Fluorescently labelled primers and dNTPs are vulnerable to degradation by exposure to light, therefore the exposure of the primer to daylight before and after incorporation were minimised wherever possible. An effort was made to prevent the oil overlaying the PCR products contaminating the samples resolved on the Genescan® gel.

The amplification of the DNA for estimation of repeat size using Genescan® software was optimised. In the first instance fluorescently labelled PCR products were generated by incorporating fluorescently labelled dNTPs during the PCR amplification. Initially a low concentration dNTP PCR buffer was used in tandem with fluorescently labelled dNTPs to increase the efficiency of the incorporation of the fluorescently labelled dNTPs. For each amplification, 1 µl of tail lysate (1 ng of DNA) in a 7 µl reaction containing 1 mM each of DM-H and DM-DR with 1X low dNTP concentration PCR buffer with and 0.175 units of *Taq* polymerase and R110, blue, fluorescently labelled dNTP's (0.017 mM). The sizing reactions were cycled 30 times, under standard PCR conditions, with an annealing temperature of 68°C. The success of the PCR amplification was checked in the first instance by resolving 3 µl of the PCR product on a 1.5 % agarose check gel (w/v). The PCR product was prepared for resolution on the Genescan® gel by mixing in a 1:1 ratio with Genescan mix. The samples

were mixed and then heated to 90°C for 2 minutes and placed on ice until they have been loaded onto the 6% acrylamide gel (v/v).

A single PCR product was expected for each *Dmt* transgene amplified whose size in base pairs and subsequently be determined. In the first instance when the products of this PCR amplification were resolved on an agarose gel, multiple bands were present (Figure 4.1.A). These bands were also seen when the PCR products were resolved on the Genescan® gel (not shown). To improve this position the concentration of fluorescently labelled dNTPs was varied (0.006-0.017M), the annealing temperature of the PCR amplification varied (68 - 74°C) and the number of cycles of PCR varied (25-30). In the first instance none of these combinations reduced the number of bands, just their intensity. It was observed that the multiple bands visible on the agarose gel were visible on the Genescan® gel and while all samples were run on agarose gels, only a small selection was resolved using Genescan® acrylamide gels, as these were a more limited resource.

Use of the alternative primer combination to amplify across the repeat, DM-C and DM-BR also resulted in multiple bands as did the replacement of the blue fluorescent dNTP mix, R110, with green fluorescent dNTP, R6G (Figure 4.1.B). However, replacement of the low dNTP concentration PCR buffer, with standard PCR buffer resulted in discrete bands for each primer combination. Comparison of the DM-H / DM-DR PCR products and the DM-C / DM-BR PCR products when generated with the green dNTPs found the dye band that was generated resolved on both the 6% and the alternative 4.75% acrylamide gel (v/v) in the same location as the PCR product, obscuring the interpretation of the results (not shown). The DM-H / DM-DR primer combination that was used widely for the SP-PCR analysis was chosen has the primer pair of preference for tail sizing. The substitution of the fluorescently labelled dNTPs with the fluorescently labelled primer DM-DRF (Figure 4.1.C) ensured that all of the PCR product generated was fluorescently labelled and eliminated the problem of the dye band distorting the PCR product size. The efficiency of the tail typing PCR amplification was further improved by reducing the concentration of the template DNA. The optimal PCR amplification conditions for the estimation of the transgenic repeat size by Genescan® software were determined to be 0.5 µl of diluted tail lysate (~0.5 ng), 0.25 Units of Bioline *Taq,* 1 X PCR buffer and DM-H and DM-DRF to a final concentration of 1µM in a 10 µl volume. If Promega Ampli *Taq* was used instead of Bioline *Taq* then the PCR reaction also contained 5% DMSO. Each reaction was cycled 28 times under standard conditions, with an annealing temperature of 68 °C. The samples were resolved on a 6% acrylamide gel using the ROX 2500 marker (Figures 4.2 and 4.3). Comparison of the same samples between a 4.75% and a 6% gel
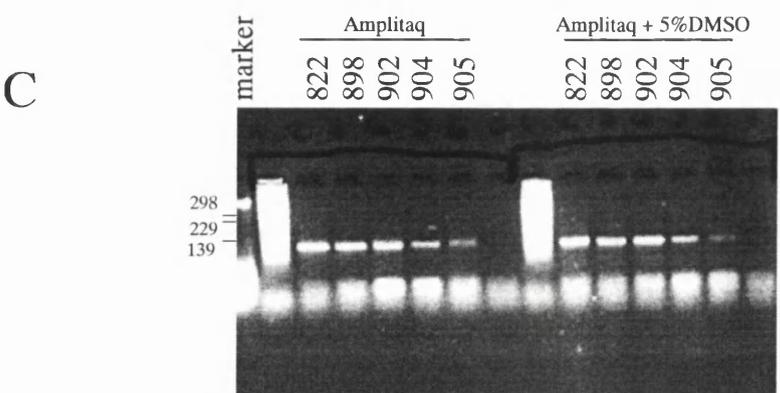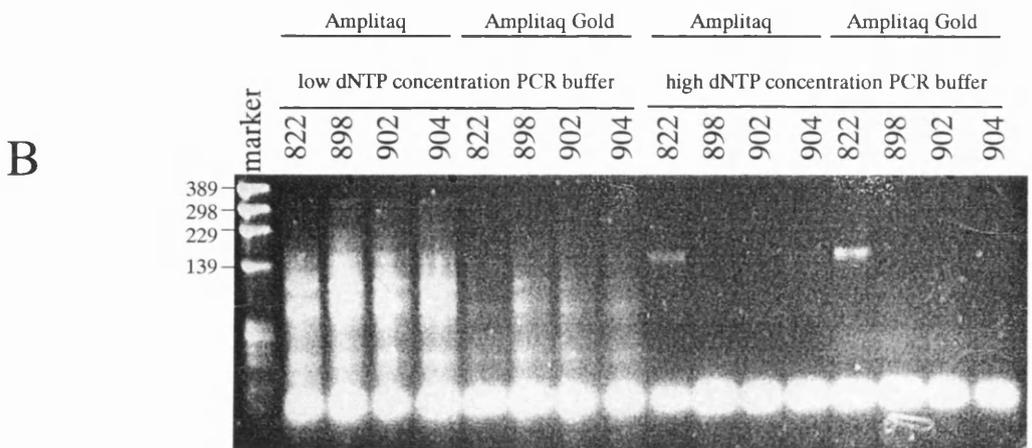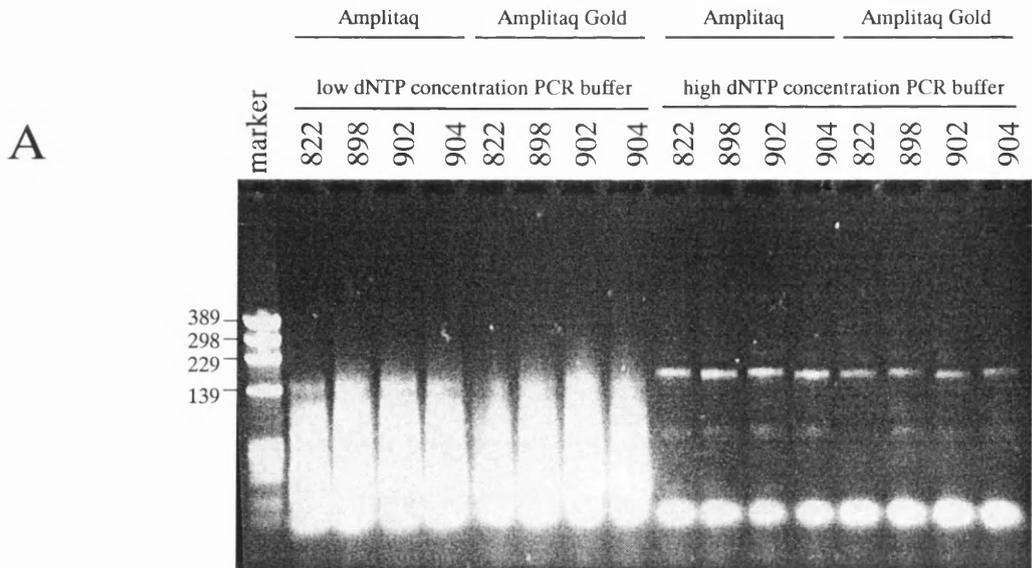
Figure 4.1
PCR amplification for Genescan analysis
Multiple PCR reactions were carried out on mouse tails to determine conditions suitable for the amplification of mouse tails DNA for analysis using Genescan software.
1 μl of tail lysate (1 ng of DNA) in a 7 μl reaction containing 1 mM each of DM-H and DM-DR with 1X PCR buffer with and 0.175 units of *Taq* polymerase and fluorescently labelled dNTPs (0.017 ul mM). The sizing reactions were cycled 30 times, under standard PCR conditions, with an annealing temperature of 68°C. The success of the PCR amplification was checked in the first instance by resolving 3 l of the PCR product on a 1.5 % agarose check gel (w/v).
(A) Shows PCR products generated with blue fluorescently labelled dNTPs (R110).
(B) Shows PCR products generated with green fluorescently labelled dNTPs (R6G).
(C) Shows PCR products generated with fluorescently labelled DM-DRF, fluorescently labelled dNTPs were not incorporated into these reactions.
The marker has been converted to triplet repeats.
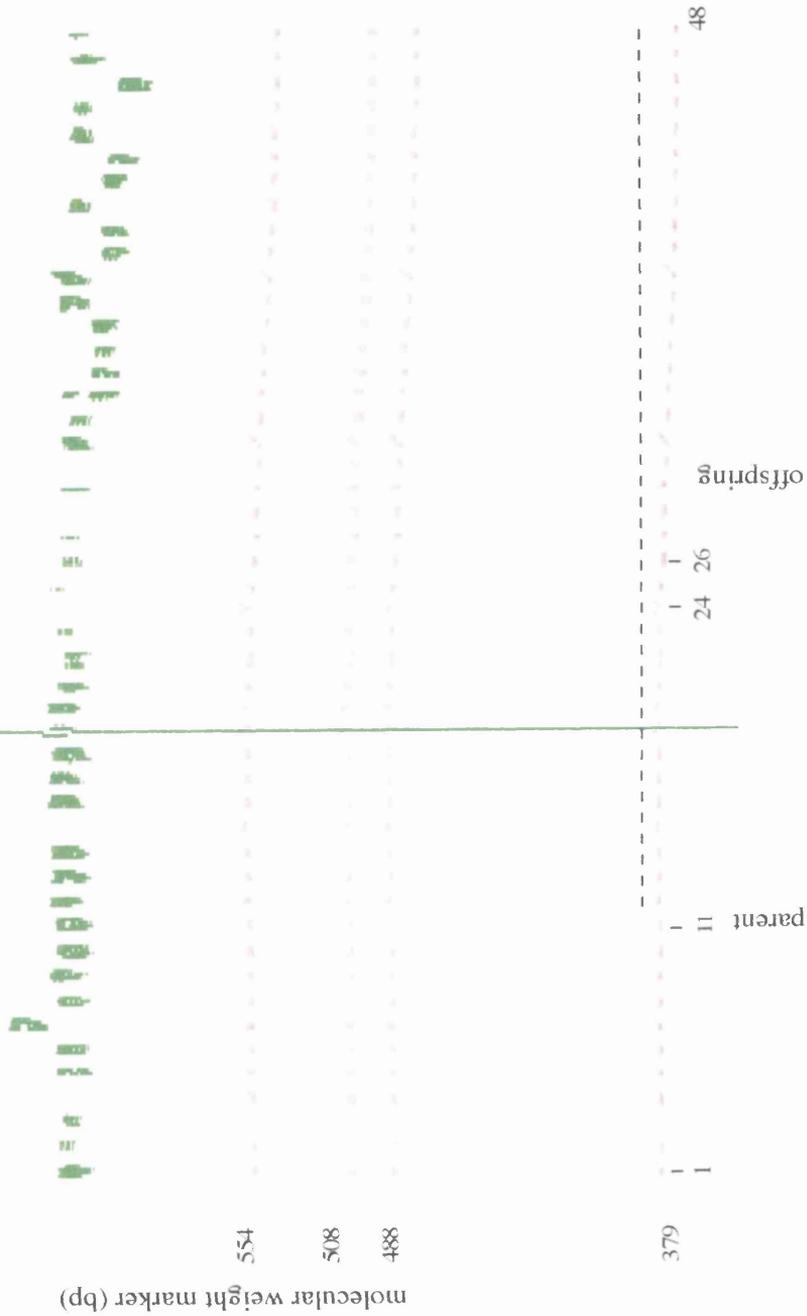
Figure 4.2
Dmt-E mouse tails resolved on a Genescan gel
Fluorescently labelled _Dmt_ PCR products amplified from the tail DNA of _Dmt_-E mice. The parent, 2654 (male) is shown in lanes 11. Indicated are lanes containing offspring of 2654, including 3254(lane 24) and 3256 (lane 26). The profiles of the indicated lanes are shown in Figure 4.3.

94
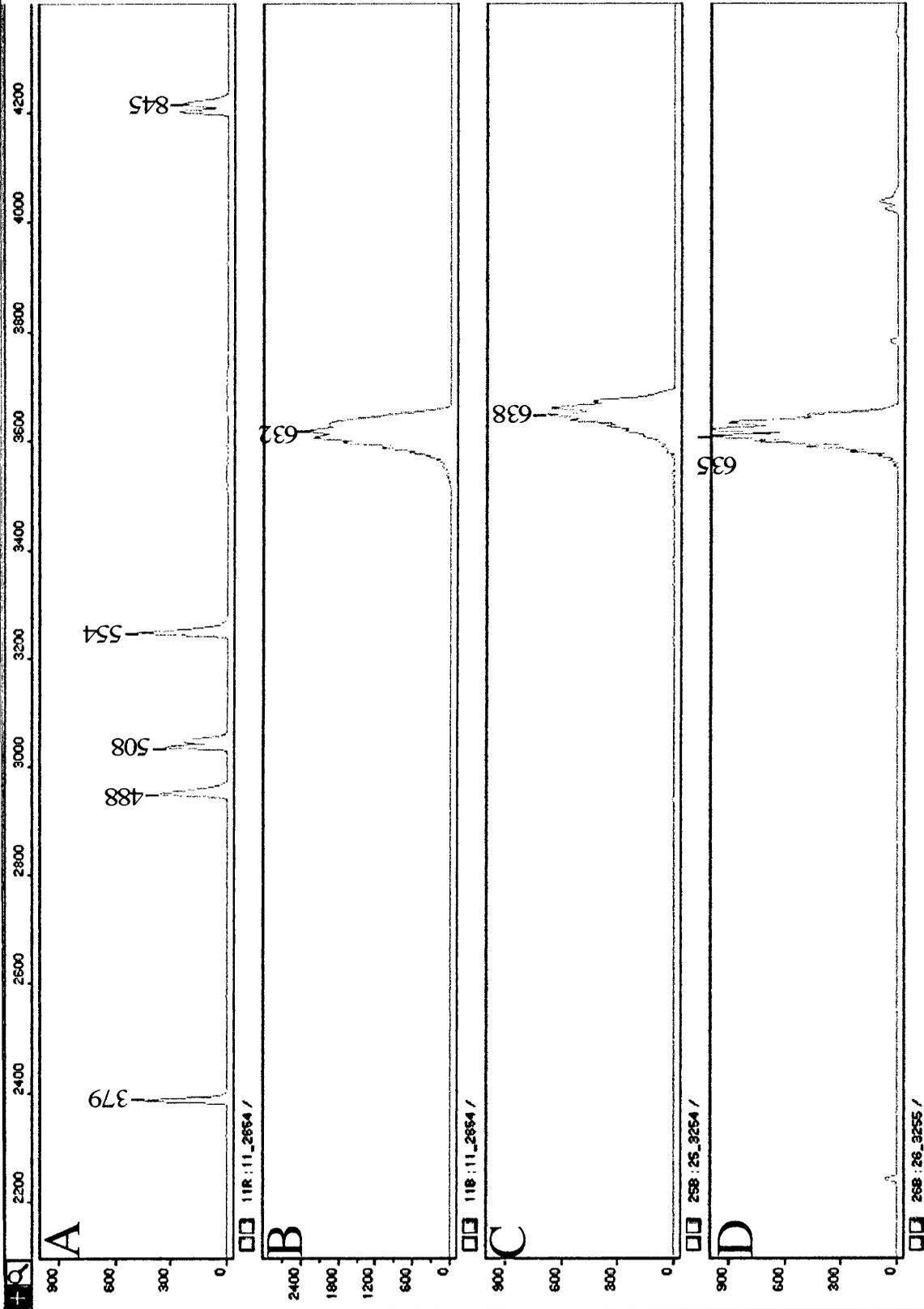
Figure 4.3
Genescan profiles of 3 *Dmt*-E mousetails
The triplet of each *Dmt*-E mouse, B: 2654, the transmitting male parent, C: 3254, and D: 3256 both male offspring were resolved against D: a ROX 2500 molecular weight marker. The molecular weight of 2654, 3254 and 3256 were determined as 631, 638 and 635 bps and 149, 151 and 150 repeats respectively.

found the repeat resolution to be clearer using a 6% gel. The estimation of the repeat size of a particular mouse varied by several base pairs when resolved on different gels despite maintenance of constant conditions. Comparison of the same sample and 3 different gels found variation of between 1 and 4 bases (data not shown), this meant that the discrepancy was not greater than one repeat. The variation was attributed to minuscule differences in the proportions of the acrylamide gel that resulted from discrepancies in the weighing and measuring of the components of the gel. This disparity was minimised by making up an acrylamide gel batch mix for 10 gels at a time thus reducing the level of variation from gel to gel. It was also decided that as the important observation was the differences in repeats of the parent when compared to the offspring the parental repeat was always resolved on the same gel as the offspring.

4.4 Transmission studies in *Dmt*-D and *Dmt*-E male and female mice

Initial observation of the transmission pattern of the *Dmt* transgene reported the frequency of transgenic transmission to be 50% for *Dmt*-D (40 transgenic mice out of 80 transmissions) and 46% (58 transgenic mice out of 126 transmissions) for *Dmt*-E (Monckton *et al.*, 1997). In contrast at the end of this study the transmission rate was 58% in *Dmt*-D (n=291) and 59% (n=208) for *Dmt*-E. Analysis of transmission using the Chi square test found both rates of transmission to be significantly different than the expected value, (p=0.5), in both cases. Combining the results from both studies the transmission rate for *Dmt*-D was 56% and *Dmt*-E was 54%.

Dmt-D	Sex of parent		
	male	female	total
Wild type	75	48	123
Dmt-D	92	76	168
total	167	124	291
% transgene	55	61	58

Table 4.1 Germline transmission of the *Dmt*-D transgene

Dmt-E	Sex of parent		
	male	female	total
Wild type	40	46	86
Dmt-E	77	45	122
total	117	91	208
% transgene	**65**	**49**	**59**

Table 4.2 Germline transmission of the *Dmt*-E transgene

Examination of the transmission rate in each line and each sex found the transmission of the transgenic repeat by *Dmt*-E males was significantly different from the expected 50%, (p=0.01) but the female transmission pattern was not. While examination of the transmission rate in *Dmt*-D males was significantly higher then expected, (p=0.9) but especially so in Dmt-D females (p=0.5).

4.4.1 Repeat transmission pattern in *Dmt*-D and *Dmt*-E

The difference in repeats between parent and offspring was determined by

Offspring repeat size - Parental repeat size = Difference in repeats

Examination of the transmission pattern of male *Dmt*-D mice (n=84), female *Dmt*-D mice (n=57) was determined (Figure 4.4.A) and male *Dmt*-E mice (n=93), female *Dmt*-E mice (n=45) was determined (Figure 4.4.B). Not all of the *Dmt*-D and *Dmt*-E mice were sized due to time constraints.

As demonstrated in Monckton *et al* (1997) the male transmissions are prone to expansions while the female transmissions are deletion prone. However, a difference in the degree of variability of the transmitted repeat is visible between the two lines, with a greater degree of variation observed in both sexes by the *Dmt*-D mice. Monckton *et al* (1997) had previously observed there to be a greater variation in the male *Dmt*-D transmissions than the male *Dmt*-E transmissions.
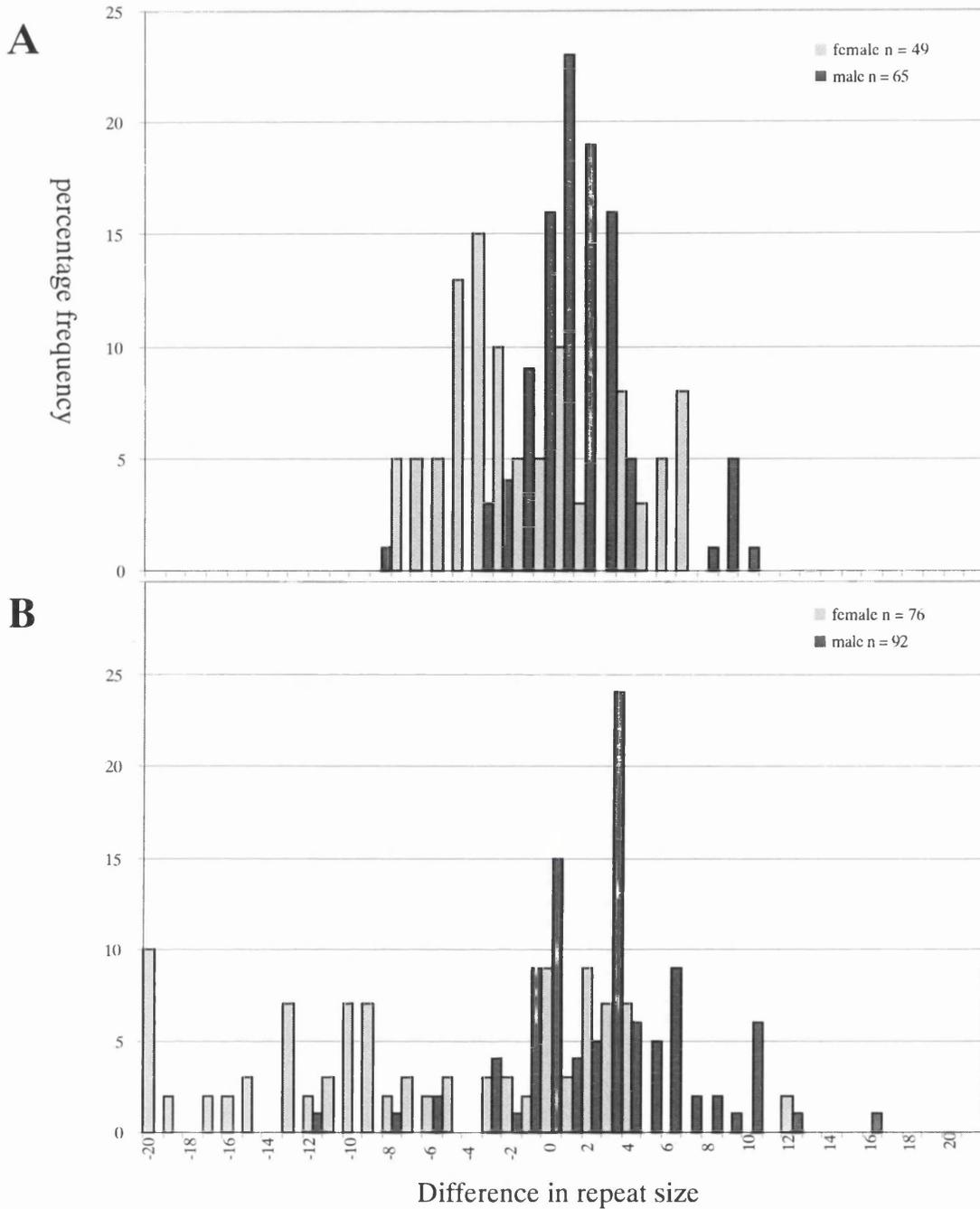
Figure 4.4
The *Dmt* repeat difference between parent and offspring
The difference between the parental transgenic repeat size and offspring transgenic repeat size is shown for male and females: A: *Dmt*-E transmissions and B: *Dmt*-D transmissions.

	deletions (%)	no change (%)	expansions (%)	mean repeat difference
Dmt-D females	63	9	28	-5.9
Dmt-D males	19	15	66	+2.7
Dmt-E females	64	10	25	-1.8
Dmt-E males	31	22	47	+0.7

Table 4.3 Rate of repeat mutation in *Dmt*–D and *Dmt*–E males and females

The parent offspring repeat difference observed in female *Dmt*-D mice is between –12 (mean deletion) and +4 (mean expansion) repeats compared to a range of –4 (mean deletion) and + 4 (mean expansion) repeats in *Dmt*-E females. The majority of the *Dmt*-E female transmissions, 64% were deletions. A similar effect was observed between the males in each line, with the mean range in repeat size in *Dmt*-D males being between –3 (mean deletion) and +6 (mean expansion) and –2 (mean deletion) and +3 (mean expansion) in *Dmt*-E males.

4.4.2 The effect of parental age and the repeat size transmitted in *Dmt* mice

To establish if the age of the parent when it transmitted the repeat to its offspring effected the size of the repeat transmitted, the age of each of the parents transmitting the transgene was determined on the day of conception for male and female *Dmt*-E and *Dmt*-D mice (Figure 4.5). Linear regression analysis revealed no significant correlation between the age of the transmitting parent and the size of repeat transmitted in either of the lines or in either of the sexes. All statistical analysis in this thesis was carried out under the supervision of Mr Grant Hogg of the University of Glasgow. Direct comparison of the difference in repeat transmission pattern and the age of transmitting parent of the *Dmt*-D and *Dmt*-E males (Figure 4.6), revealed little variation in the *Dmt*-E mice, in *Dmt*-D however, although large parent offspring differences occur at all ages, some of the largest variations were observed in the older *Dmt*-D males. A similar comparison of the difference in repeat transmission pattern and the age of transmitting parent of the *Dmt*-D and *Dmt*-E females (Figure 4.6) also revealed little variety in the repeat transmitted despite the age of the parent in *Dmt*-E females, while most of the larger variations in repeat size in *Dmt*-D females were transmitted by older mice.

It has been suggested that the sex of the conceived mouse has an effect on the repeat size presented by that mouse (Kovton *et al.*, 2000). To examine whether sex of the offspring has
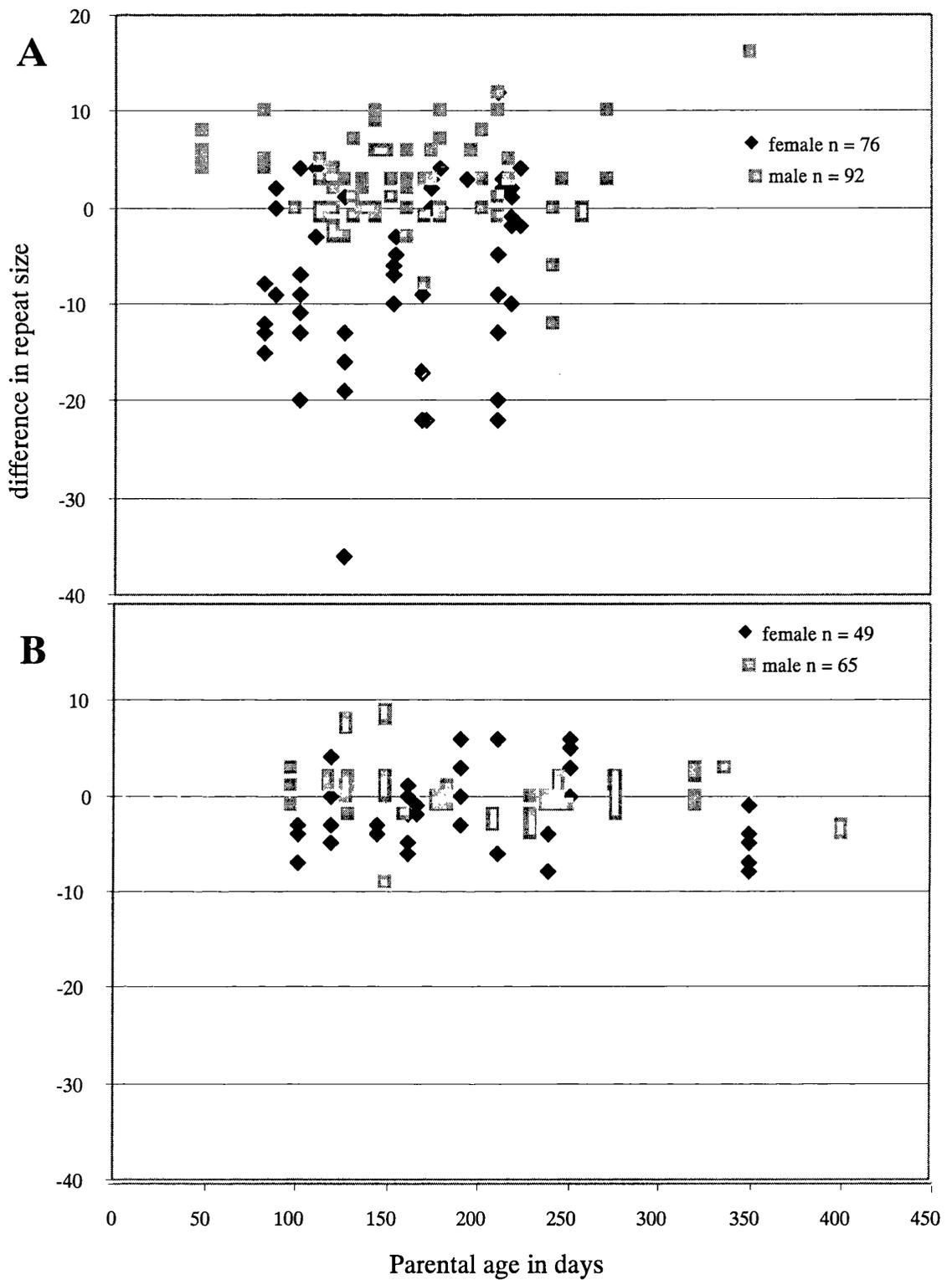
Figure 4.5
Age of parent against differences in repeat size transmitted
The difference between the *Dmt* repeat size between each parent and its offspring is plotted against the age of the parent (days) at the time the offspring was conceived for A: *Dmt*-D male and female mice and B: *Dmt*-E male and female mice.
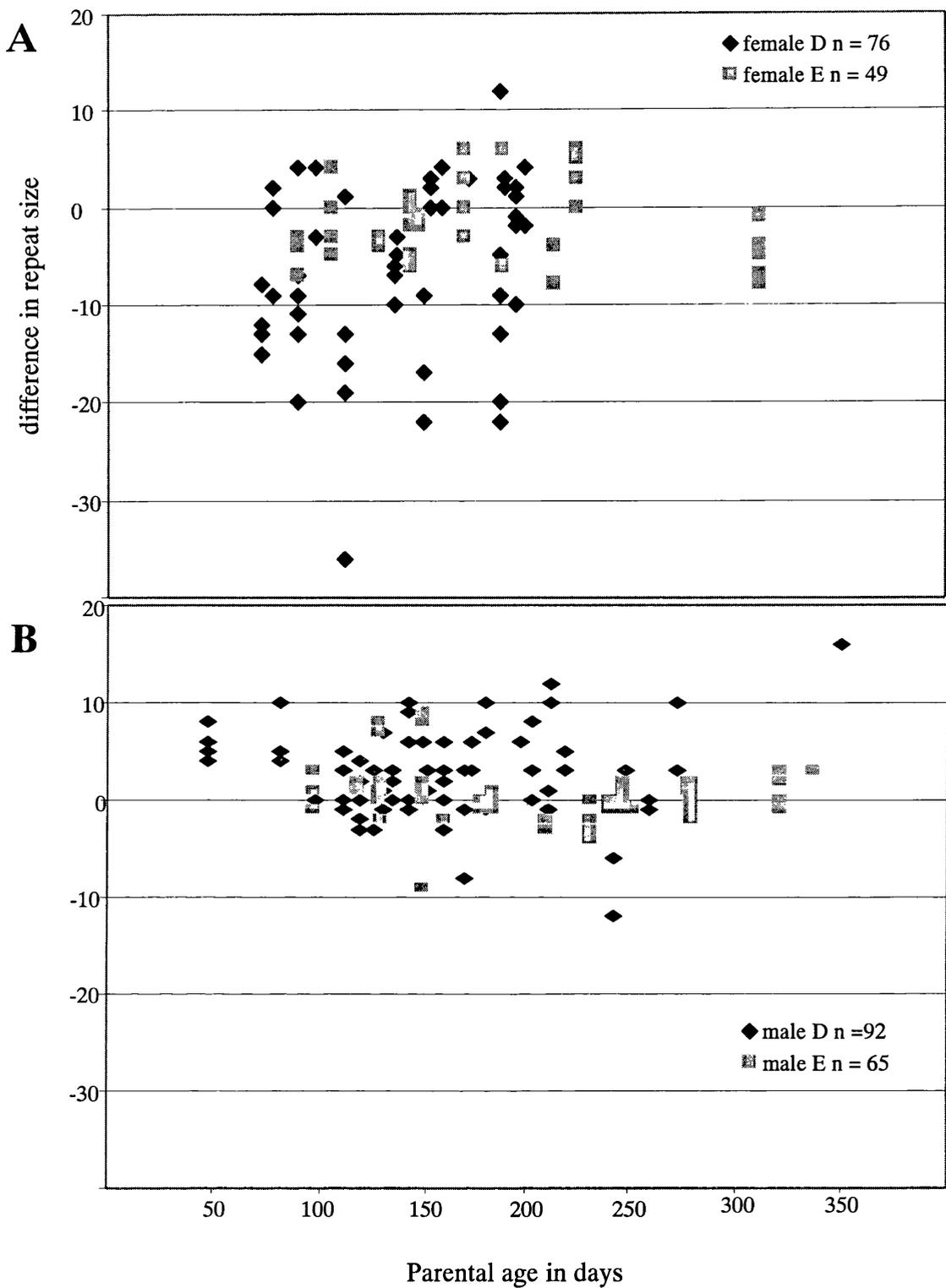
Figure 4.6
Comparison of repeat transmission pattern between *Dmt*-D and *Dmt*-E mice
The difference between the *Dmt* repeat size between each parent and its offspring is plotted against the age of the parent (days), at the time the offspring was conceived for A: female*Dmt*-D and *Dmt*-E mice and B; male *Dmt*-D and *Dmt*-E mice.

influenced the parent offspring difference, each of the four sets of data, the differences in repeat size for the *Dmt*-D male and female parents, and the *Dmt*-E male and female parents were themselves split into male and female groups (Figure 4.7 and Figure 4.8). Although the sex of the offspring made no apparent difference to the *Dmt*-D transmissions or the female *Dmt*-E transmissions, the size of repeat carried by the offspring of male *Dmt*-E mice split into two distinct modes with the females carrying repeats closer to that of the transmitting parent than the males. Application of the Kolmogorov Smirnov test to each of these modes found that the 2 modes were significantly separate.

4.5 Examination of triplet repeat instability in the testes and sperm of *Dmt*-D mice at different time points

To investigate whether the instability observed in the tissues of *Dmt*-D mice was also visible in the germline of male mice, DNA was prepared from the sperm and testes of *Dmt*-D mice at 2 and 13 months. The DNA was diluted to 1200 pgμl^{-1}, 120 pgμl^{-1}, 12 pgμl^{-1} in dilution buffer containing 0.1μM DM-H. The 1200 pgμl^{-1} and 12 pgμl^{-1} concentrations were used as templates for multiple SP-PCR amplifications (Figure 4.9). The sperm DNA amplified poorly, while level of repeat variation in the testes DNA increases with age.

4.6 Amplification of single oocyte by PCR

To allow assessment of the repeat stability level in the unfertilised oocyte when compared to the repeat carried by the transgenic offspring, it was proposed that *Dmt* transgenic females be superovulated using PMSG and HCG, before mating with a vasectomised male (Methods 2.2.5.4) to bring a maximum number of ova to a common mature state before they were harvested by flushing from the oviduct into M2 media. Single ova were manipulated using a mouth pipette into 200 μl thin walled PCR tubes. The presence of each ova, which was isolated in ~ 10 μl of M2 media, was confirmed by checking the contents of each tube underneath the microscope and ~ 20 μl of mineral oil placed over the contents of each tube to prevent loss of contents by evaporation. The ova were stored on ice until the harvest was completed. From this point onwards, all of the manipulation of the ova was carried out in a laminar flow hood.
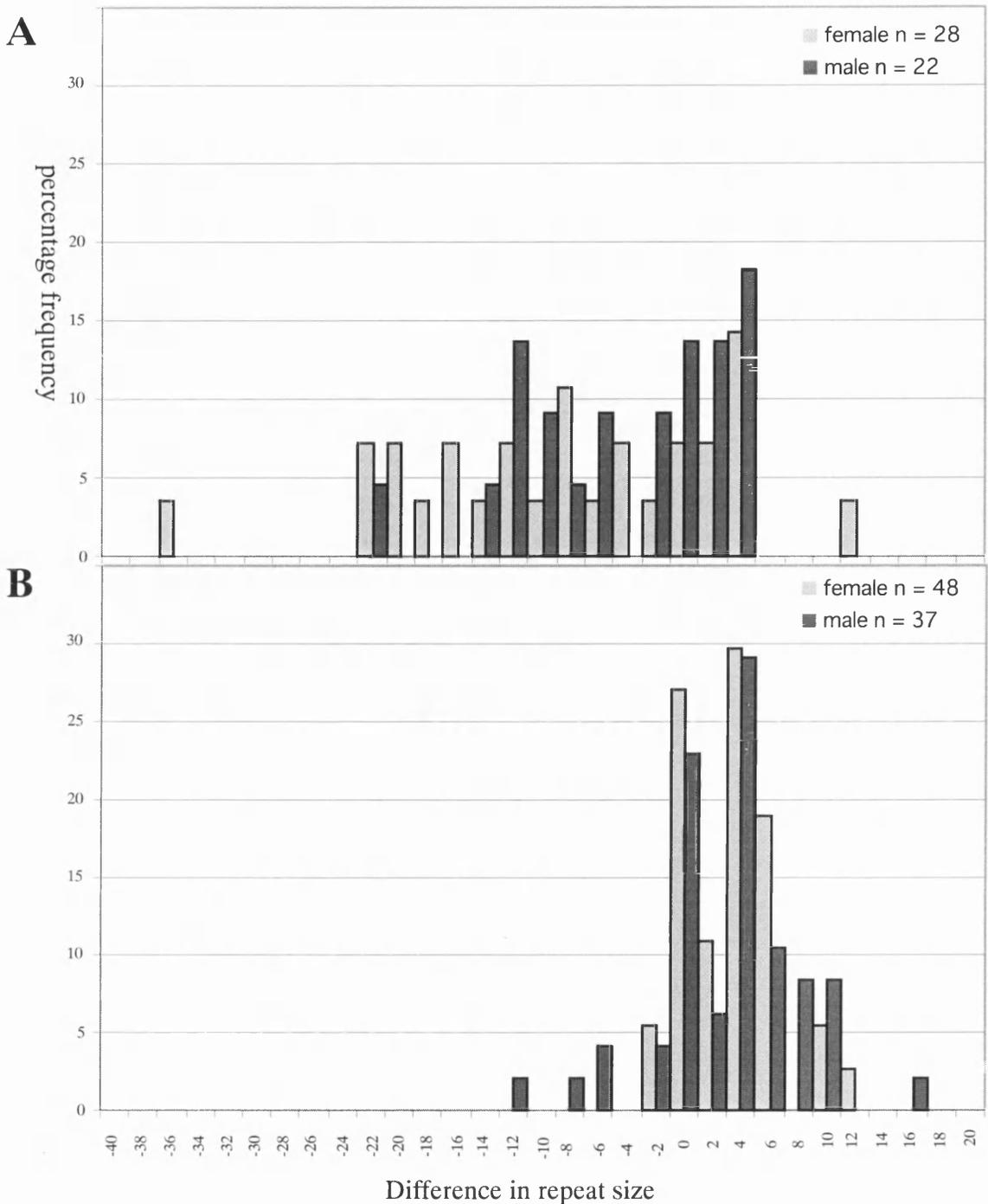
Figure 4.7
The sex distibution of the parent offspring difference in *Dmt*-D repeat transmissions
The distribution of the *Dmt*-D parent offspring repeat difference in the offspring of A:
Dmt-D females and B; *Dmt*-D males.
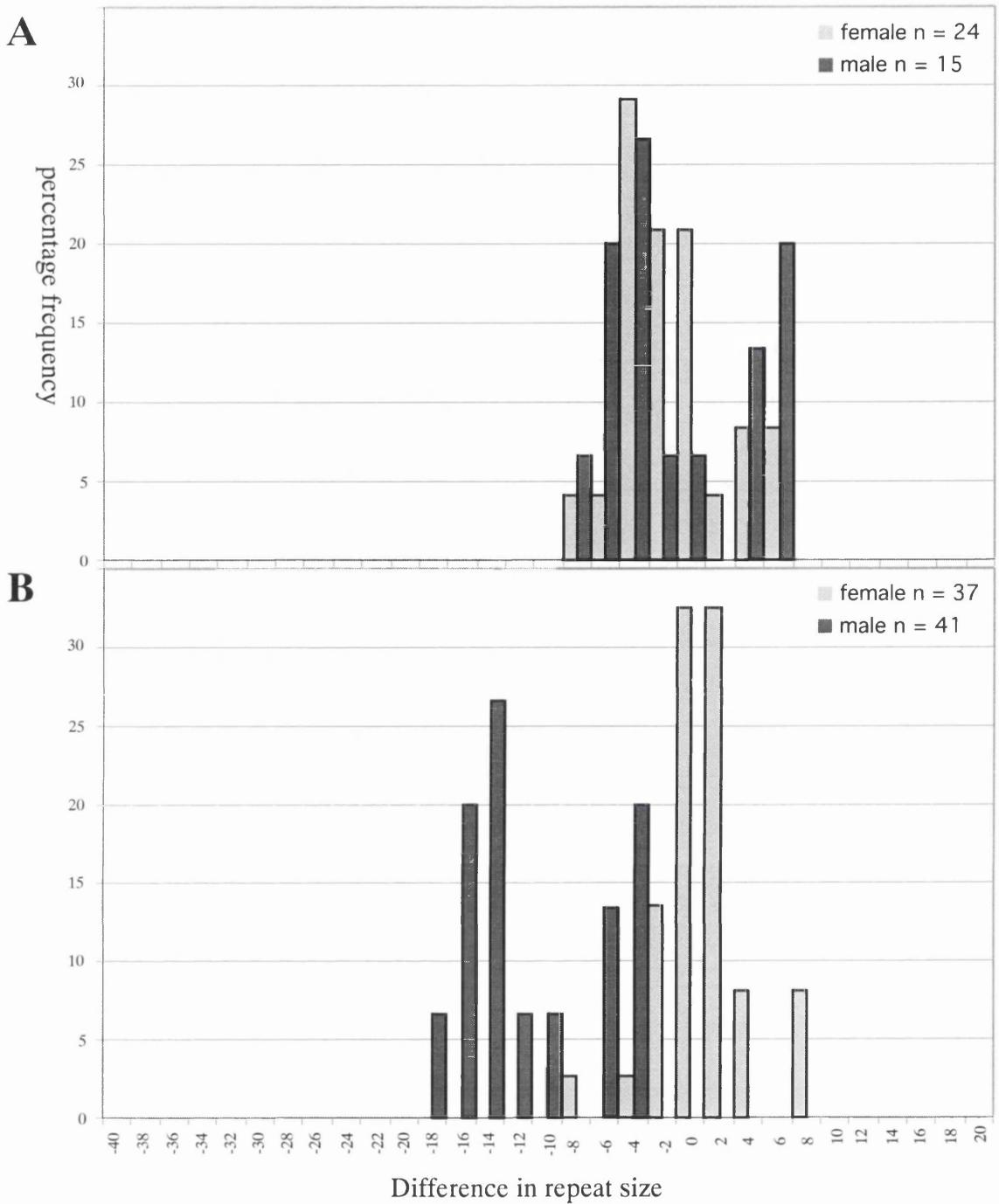
Figure 4.8
The sex distibution of the parent offspring difference in *Dmt*-E repeat transmissions
The distribution of the *Dmt*-E parent offspring repeat difference in the offspring of A: *Dmt*-E females and B; *Dmt*-E males.
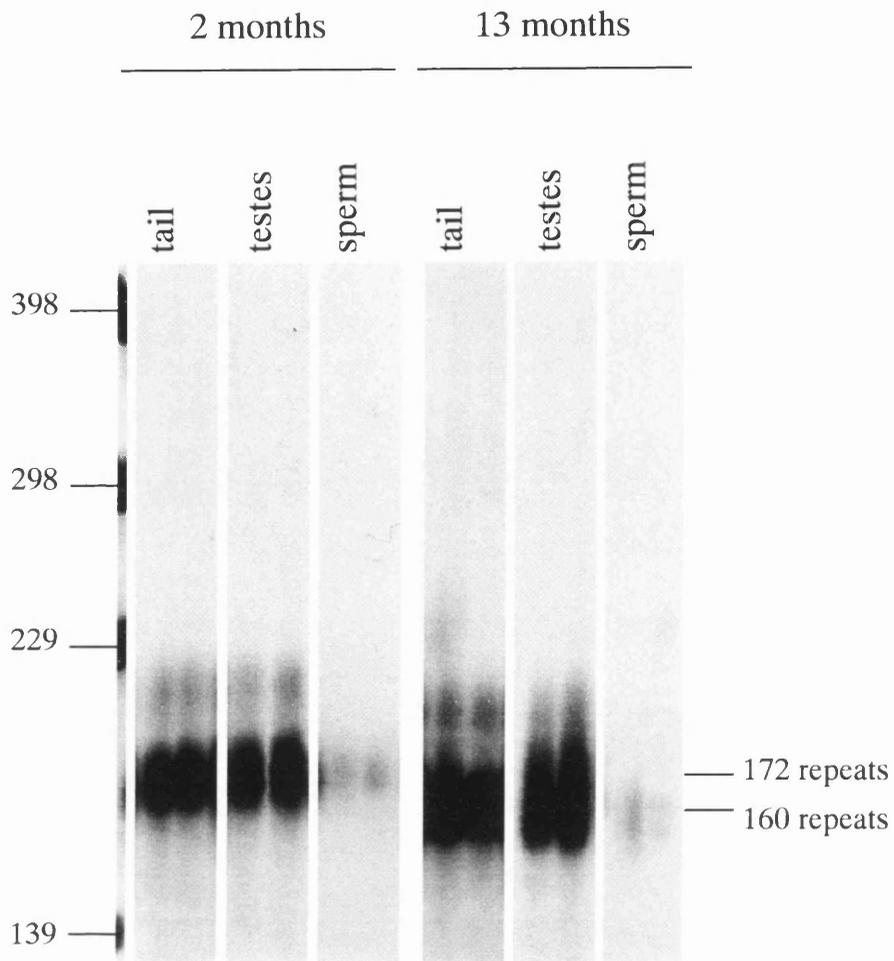
Figure 4.9
Small pool PCR amplification of the sperm and testes of _Dmt_-D mice
Between ~10 and 100 molecules from the mature tail, testes and sperm of a 2month old
Dmt-D mouse with a repeat size of 172, and a 13 month old _Dmt_-D mouse with a repeat
size of 160 were used as the template for duplicate small pool PCR amplifications which
were resolved on a 1.25% (w/v) agarose gel before transfer to a nylon membrane and
Southern hybridisation with DM56. The molecular weight marker has been converted to
indicate the number of triplet repeats.

Each ovum was lysed using a modified method based on that described in, Kaytor *et al.* (1997). 5μl KOH lysis buffer was overlaid onto the oil in each tube and the contents mixed by gentle tapping of the tube, incubated at 65 °C for 10 minutes before neutralising by adding 5.2μl of 200mM Tricene. In the first attempt only the transgenic repeat was amplified using DM-H and DM-BR. Assuming that the final volume for each lysed oocyte was ~ 20 μl, a 40 μl reaction was set up using 1X PCR buffer, 0.035 Units of *Taq* and 0.2 μM each of DM-H and DM-BR. Each PCR amplification was cycled 30 times under standard conditions with an annealing temperature of 68 °C. However, the failure of the oocyte to act as template for the PCR amplification (possibly due to failure of oocyte lysis) could not be discerned from the failure to amplify the repeat. This was because the oocyte did not contain the transgene as all of the superovulated mice were hemizygous for the transgene. In subsequent attempts, the lysed oocytes were used as the template for PCR to amplify the mouse *Usf2* gene, using mUSF-A and mUSF-BR and the transgenic repeat, using DM-H and DM-BR in the same PCR amplification. As all of the oocytes should contain the mouse *Usf2* gene, the presence of this PCR product should confirm that the oocyte had indeed been successfully lysed and amplified from. Assuming that the final volume for each lysed oocyte was ~ 20 μl, a 40 μl reaction was set up using 1X PCR buffer, 0.035 Units of *Taq* and 0.2 μM each of DM-H, DM-BR, mUSF-A and mUSF-BR. Each PCR amplification was cycled 30 times under standard conditions with an annealing temperature of 63 °C. Initially 1 μl of this PCR product was used as a template to amplify the transgenic repeat using 1X PCR buffer, 0.035 Units of *Taq* and 0.2 μM each of DM-H and DM-DR in a 7μl volume for 30 cycles under standard PCR reactions with an annealing temperature 68 °C. If the mouse *Usf2* primer pair had also been used then a further 1 μl of first round of PCR amplification was used as a template to amplify the mouse *Usf2* product using 1X PCR buffer, 0.035 Units of *Taq* and 0.2 μM each of mUSF-A and mUSF-BR in a 7μl volume for 30 cycles under standard PCR reactions with an annealing temperature 63 °C. Amplifying the PCR products a second time did not identify the presence of any transgenic oocytes that were not identified in the first round.

All of the PCR products, from both rounds of amplification were resolved on a 1.5% agarose gel before transfer to nylon membrane and Southern hybridisation with an appropriate probe. The first round of the PCR amplification was kept to use as the template for subsequent Genescan® reactions to size the repeat in transgenic oocytes. It was decided that the second amplifications were not necessary.

4.6.1 Amplification of a single oocyte from *Dmt*-D and *Dmt*-E by PCR

In the first instance a single *Dmt*-D mouse, aged 15 weeks was superovulated and of the four oocytes present, 3 were isolated, lysed and subjected to two rounds of PCR amplification. Of the 3 oocytes, 1 of the 3 resulted in the amplification of the transgenic repeat. Having discerned in the first instance that the transgenic repeat could be isolated in this manner further mice were subjected to superovulation and PCR amplification in the presence of mUSF-A / mUSF-BR. The superovulation experiments are summarised in table 4.2 (Figure 4.10). The yield of oocytes from the superovulated mice was very poor when compared to that obtained from other mice in the same facility, where as many as 40 oocytes were isolated from a single female. In the first instance (attempts 1 and 2) the low yield was attributed to the failure of the PMSG to stimulate oocyte production, it was observed in at least 3 of the mice that one of the oviducts yielded nothing at all, which was observed to be unusual. An alternative source of PMSG was used for attempt 3, however, the number of isolated oocytes continued to be low. This could have been accounted for by the age of the mice as mice aged less than 10 weeks are recommended for the isolation of embryos for cryopreservation. However, the nature of the question addressed by this experiment requires that older mice be studied. It is possible that the *Dmt*-D and -E transgenic lines do by their nature yield fewer oocytes than wildtype FVB/n. However, this observation has not been made for *Dmt*-D and –E or indeed *Dmt*-C, although the *Dmt*-B litter sizes dropped to half that expected. Attempts to lyse the isolated oocytes and ascertain the presence of the transgene were unsuccessful at subsequent attempts. This was possibly due to the failure of lysis of the oocyte, in each case. It was suggested that the presence of the M2 buffer could be inhibiting the lysis or alternatively the efficiency of the tricene buffer to neutralise the solution, which in turn resulted in the inhibition of the PCR amplification. It was also around this time that the mUSF-A primer batch in use failed to work. If this was responsible for the lack of mouse *usf2* DNA amplified, the transgenic repeat primer pair may still have worked. However, it was thought necessary to ascertain the number of oocytes containing the transgene accurately, as well as the number of repeats present.
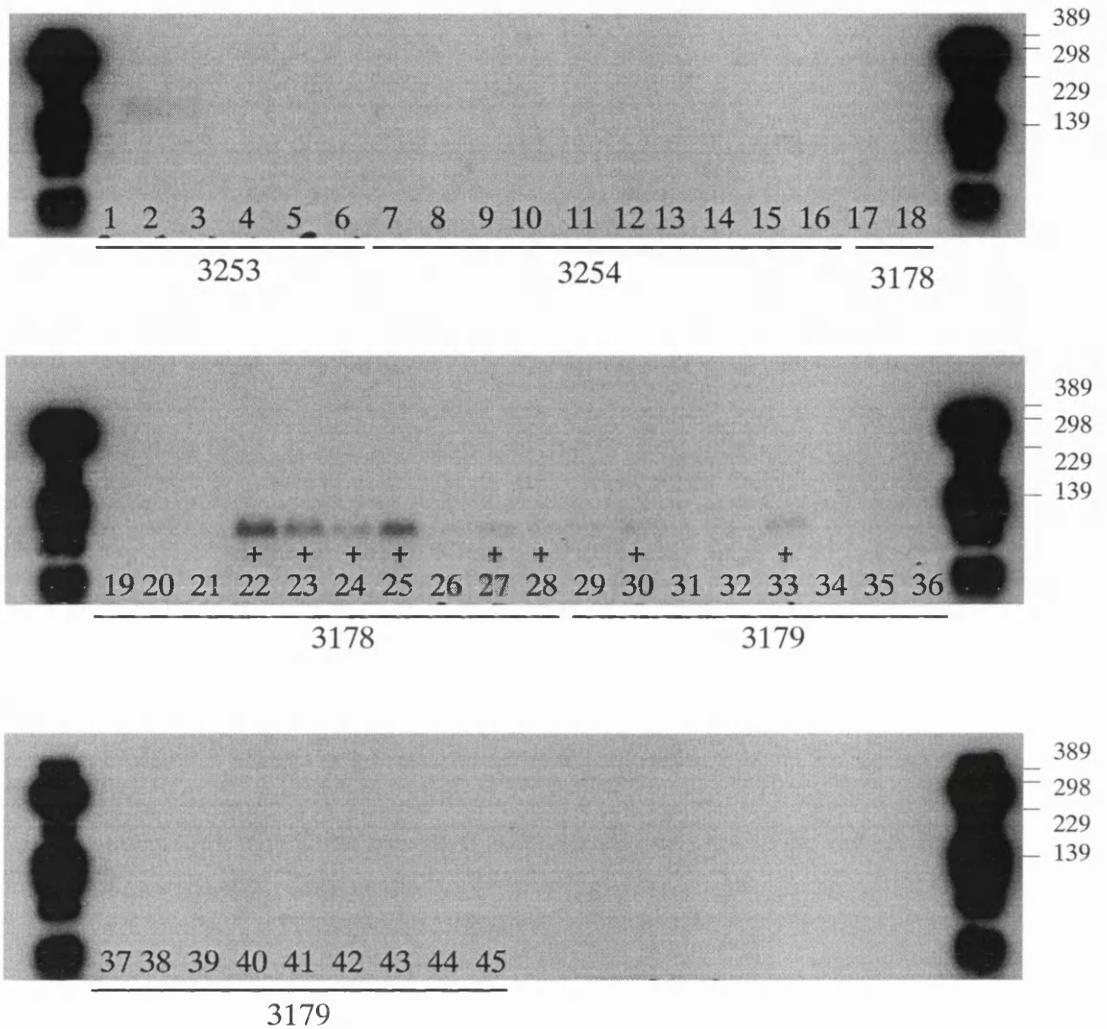
Figure 4.10
Oocytes isolated from *Dmt* mice after superovulation
After superovulation, individual oocytes wereisolated and lysed before being used as templates for SP-PCR amplification. 1X PCR buffer, 0.035 Units of *Taq* and 0.2 μM each of DM-H and DM-DR in a 7μl volume for 30 cycles under standard PCR conditions with an annealing temperature 63•C. The PCR products are resolved on a 1.5% gel before transfer to a nylon membrane and hybridisation with DM56.

Attempt	*Dmt* line	Age of mouse super ovulated (weeks)	Oocytes isolated	Transgenic repeat amplified
1	*Dmt*-D	(3110) 15 weeks	4	1
2	*Dmt*-D	(3178) 11 weeks	12	6
2	*Dmt*-D	(3179) 11 weeks	17	2
2	*Dmt*-E	(3253) 15 weeks	6	none identified
2	*Dmt*-E	(3254) 15 weeks	10	none identified
3	*Dmt*-D	(3107) 23 weeks	2	none identified
3	*Dmt*-E	(3051) 25 weeks	18	none identified

Table 4.2 Summary of oocytes isolated from *Dmt* mice after superovulation

The isolation of oocytes from the *Dmt* strains proved inefficient and inconclusive. The isolation of gametes from *Dmt* mice on alternative strain background may prove more successful.

4.7 Discussion

The observation of sex differences in repeat transmission patterns of the *Dmt* lines is in common with observations made in triplet repeat diseases although these patterns are disorder dependent. DM1 females carrying large repeats transmit larger repeats to their offspring. Maternal DM1 transmissions account for the majority of CDM cases (Ashizawa *et al.*, 1994; Tachi *et al.*, 1997) suggesting that location of repeat does effect its repeat transmission instability profile. The degree of this variation differs between *Dmt*-D and *Dmt*-E, transgenes that have different integration sites and this highlights the role that must be played by *cis* acting factors. Seznec *et al* (2000) observe the transmission patterns of several DM1

transgenic lines which contain ~300 repeats. The DM 300-328, DM 300-1112 and DM 300-1177 lines were generated from a transgene which encodes the expanded DM1 repeat within ~ 45 kb of flanking genomic sequence including the coding sequence for *DMPK*, *DMWD* and, *SIX5*. Observation of the differences between the repeat transmitted between parent and offspring for each line observed DM 300-328, DM 300-1112 and DM 300-1177 transmitted 86.5%, 88% and 95.5% expansions respectively. The effect of the transmitting parent sex on the repeat size was not reported. Each of the different lines presented different ranges of repeat events, the range for DM 300-328 was between +1 and +60, for DM 300-1112, between +3 and +10 and for DM 300-1177, +2 and +42. This suggests that despite the influence of the genomic DM1 DNA on the repeat, variations in the integration sites of the lines also effects the repeat instability profile presented.

The effect of the age of the parent at the conception of its offspring on the parent offspring repeat difference was also examined by Seznec *et al* (2000). They observed a positive correlation between the increasing age of transmitting fathers and the increase in the repeat size transmitted to their offspring. No correlation was reported in maternal transmissions. The absence of correlation between the age of the transmitting parent and the size of the repeat transmitted in either of the *Dmt* lines may be due to the nature of the *Dmt* transgene. The lack of DM1 flanking DNA may effect this in the same way as differences in the integration site effect the size of the transmitted repeat. As an age / repeat size effect has been observed in DM 300 mice it is less likely that the lack of correlation is due to possible differences in gamete development in humans and mice. Age of parent repeat effects have also been reported for mouse models (Kaytor, 1997; Lorenzetti *et al.*, 2000). Interestingly, the splitting of the four groups of transmitted data, that is maternal and paternal *Dmt*-E and *Dmt*-D transmissions into subsets consisting of the sex of the offspring reveal no obvious pattern in the maternal transmissions of either line. However, the paternal transmissions of the *Dmt*-E mice clearly splits into two distinct sex dependent modes. Kovten *et al* (2000) report a similar effect in male transmissions of (R6/1) Huntington's model mice, although no data is given about female transmission patterns. They suggest that the intergenerational repeat differences observed are due at least in part to events occurring within the zygote. This may be possible, though if this was happening in *Dmt*-E mice it might be expected that more than one transgenic progenitor repeat size would be present in at least some of the tissues. Perhaps the sex of the sperm cell after division but preimplantation is influencing the repeat in some way. *Trans* acting factors associated with the Y chromosome but not the X chromosome or vice

versa may effect the manner in which misalignments of the repeat are dealt with. In all other ways the mice are on a uniform genetic background.

Examination of instability in the sperm of *Dmt*-D mice was uninformative, although it was observed that the testes became more unstable with age. Pedigree data revealed no correlation between the increasing age of the transmitting father and the size of the repeat transmitted, although most of the large parental offspring repeat differences were observed in older parents. It is possible that there may be a correlation, but that the number of mice studied must be increased.

That segregation distortion was observed in *Dmt*-D and *Dmt*-E transmissions was in line with observations made of DM1 transmissions (Carey *et al.*, 1994; Hurst *et al.*, 1995). More interesting however, was the influence of the sex of the transmitting parent on the segregation distortion observed, with 65% of male transmissions resulting in offspring carrying the transgene while female transmissions of the transgene were not significantly distorted. While male and female *Dmt*-E transmissions were significantly more likely to result in offspring carrying the transgene, females were more likely to transmit the transgene than males. Again these differences between the lines may be in some way accounted for by the differences of the *Dmt*-D and *Dmt*-E integration site.

Chapter 5

Investigation of transcription levels of the *Dmt* transgenes

5.1 Introduction

It has been demonstrated in Chapter 3 that the *Dmt* transgene exhibits variation in the somatic stability of the transgenic repeat depending on the line. As the *Dmt* transgenic lines contain similar numbers of repeats and the effect of the flanking transgenic DNA is thought to be minimal, the differences observed in the four *Dmt* lines must be due to differences in the integration of the transgenes into the mouse genome. That is, that the repeat in each line is subjected to different *cis* acting factors and these differences might explain the different stability levels observed in each transgenic line.

It may be possible that the instability mechanism effecting repeat stability might involve the transcription process (Figure 5.1). If the transgene was transcribed, then the paired triplet repeats are disrupted to allow transcription of the region. When the strands reanneal it is possible that the repeats fail to anneal "in register", resulting in unpaired loops of repeats. If the DNA repair system was to recognise such "loop outs" as aberrant and "repair" them it could result in a difference in repeat length? Indeed if such a process resulted in expansions, the more often this annealing / reannealing occurred, the more likely it would be that the repeat would increase in size. As it appears that the larger a repeat is the more vulnerable it is to expansion, possibly due to the increased likelihood that the larger an expanded repeat tract the greater the chances that the repeat strands realign out of register. Therefore it could be that the degree to which the repeat region is transcribed could correlate with the level of instability observed in that tissue.

It has been observed that *DMPK* is expressed in a tissue specific manner (Erikkson *et al.*, 2000), meaning that the repeat is transcribed more frequently in some tissues. Could the differences observed in the stability of the transgenic repeat in the different lines be explained by the transcription of, or lack of transcription of the repeat? Indeed could the instability observed in *Dmt*-D be because it has integrated downstream of a promoter region and is transcribed frequently, while *Dmt*-B, -C and -E have not and are not? Speculation as to the nature of the integration site in *Dmt*-D is that it is in some way

repeat tract

paired DNA strand

Paired DNA strands separate
to allow transcription

After transcription the reannealling
repeats align out of register

The resulting loopouts are recognised
as aberrant by the mismatch repair complex

resulting in repeat deletion   no change   resulting in repeat expansion

The more often the repeat region is
transcribed, the larger the repeat
expansion will become

Figure 5.1
How the transcription of an unstable triplet repeat tract could result in expansion
The figure demonstrates how a functioning MMR complex could result in a change in the repeat
length whilst the transgene is being transcribed. Indicated are the mismatch repair heterodimers as
outlined in figure 3.1.

incorporated into a kidney related gene and that there would be some correlation of transcription levels of the transgene and nature of instability observed in *Dmt*-D.

An alternative hypothesis suggested that more than one transgene had integrated downstream of a promoter region and was therefore transcribed but differences in stability levels could be influenced by the orientation of the integrated CTG·CAG transgene. To investigate these theories a series of experiments based on reverse transcription PCR were carried out.

5.2 Transcription of the *Dmt* transgene

5.2.1 Transgenic transcription in *Dmt*-B, -C, -D, and -E

To determine whether the *Dmt* transgene was being transcribed in any of the lines, total RNA, that is RNA from the lysed nucleus as well as the cytoplasm was prepared from the heart, brain and kidney of mice from each line and from a wild type FVB/n mouse. The mice were aged between 2 and 4 months, with the exception of the *Dmt*-B mouse, which was 20 months. 3.0 µg of RNA in 5 X RNA loading buffer was electrophoresed on a 1% agarose, 0.1 % SDS gel in 1 X TAE (w/v) and stained by shaking in ~ 0.15 ng ml $^{-1}$ EtBr in distilled water. This allowed assessment of the quality of the RNA and the accuracy of the estimations of the concentration (Figure 5.2). If the RNA was of poor quality, the extraction was repeated where possible, with the exception of *Dmt*-B, where the source of archived tissue was limited. If the estimation of the concentration was inefficient the readjustment of the concentration was made by using the Kodak 1D sizing programme to estimate the concentration of the bands relative to each other.

For each sample, cDNA was generated using 2.5 µg of each RNA as a template to make cDNA with 50 ng of random prime hexamers (Methods, 2.2.7.2). The reaction that was carried out with Superscript II reverse transcriptase was referred to as (+RT). A control reaction without reverse transcriptase, (-RT) was carried out for each tissue to check for genomic contamination. There should be only RNA in the -RT control. When this was used as template for a PCR amplification there should be not be a resulting product. If however, there is genomic DNA present in the template RNA this will be amplified and the resulting band will be indicative of genomic contamination.
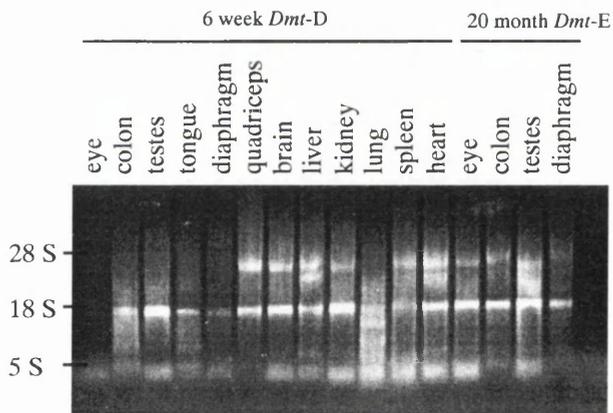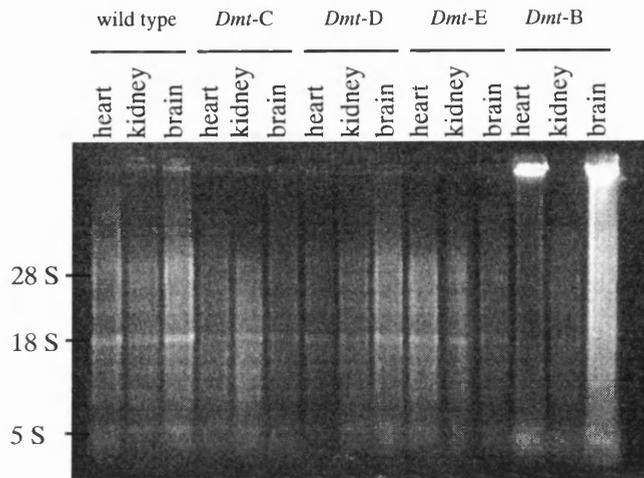
Figure 5.2
RNA check gels
Each lane contains 2.5 g of RNA as determined by spectrophotometer on a
1% agarose gel, 0.1% SDS in 1XTAE stained by shaking in distilled water
iwith ~ 0.15 ng ml $^{-1}$ EtBr w/v for 1 hour.The location of ribosomal RNA
bands is indicated.

Initially the PCR amplification of the mouse *Gapdh* product was used to determine the success of the reverse transcriptase reaction and to identify the presence of genomic DNA contamination in the RNA template (Figure 5.3). Controls to check for contamination were carried out for each set of PCR amplifications. This meant a -RT reaction for each sample. Both the +RT and -RT products were amplified with the mouse *Gapdh* primers, GAPDHA and GAPDHBR under standard PCR conditions with an annealing temperature of 68°C, which would amplify a product size of 696 bp. If genomic contamination was identified the RNA template was subjected to further DNase treatment (Methods, 2.2.9.1) until genomic contamination could no longer be detected. If cDNA could not be amplified using the mouse *Gapdh* primers, the +/- RT reactions were repeated for that RNA sample.

It is known that triplet repeats can be difficult templates for enzymes to interact with, as enzymes are known to "stall" as they progress through repeat tracts (Parsons *et al.*, 1998) and that reverse transcriptase needs time to read the template. The contact time of the reverse transcriptase enzyme with the repeat tract during incubation was increased from the 50 minutes recommended in standard protocols to 90 minutes in each case.

PCR was carried out on the RT +/- samples using CLW2F / CLW4R forward and reverse primers located in exon 2 and exon 4 of the mouse *Dmpk* gene respectively. The PCR amplifications were carried out under standard PCR conditions with an annealing temperature of 61°C. This primarily revealed a genomic transcript, 563 bp in size indicating the presence of heteronucleic RNA and spliced product, 232 bp in size (Figure 5.3).

RT-PCR was then carried out on the RT +/- samples using transgenic primers. DM-H / DM-DR flank the repeat and result in a product size dependent on the repeat number carried by the *Dmt* transgenic mouse and DM-R / DM-PRENK, which amplify a non repeat region of the *Dmt* transgene and results in a product 276 bp in size (Figure 5.4.A). These PCR amplifications were carried out under standard PCR conditions with an annealing temperature of 68°C. The products from each RT-PCR amplification were resolved on a 1.5% agarose gel (w/v) and a photograph taken before the PCR products were transferred via squash blotting (2.2.3.6) to a nylon membrane. The DM-H / DM-DR product was probed with DM-56, (Figure 5.4A)

116
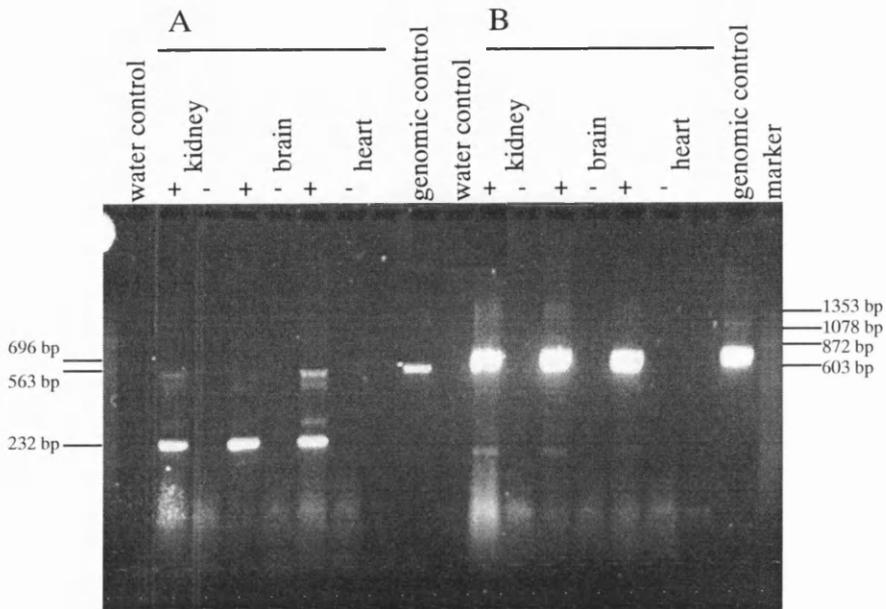
Figure 5.3
Amplification of cDNA generated with random prime hexamers amplified with
A. CLW2F / CLW4R
B.GAPDHA/GAPDHBR
2.5 µg of total RNA from each tissue was used as a template for the generation of cDNA, which was primed using 50ng of random prime hexamers. The presence of the reverse transcriptase enzyme in the reaction is indicated by + (with), or - (without). 1µl of cDNA product was used as the template for RT-PCR amplification using either A: CLW2 / CLW4 primers located in exon 2 and exon 4 of *dmpk* respectively or B: GAPDHA / GAPDHBR which amplifies a 696 bp *gapdh* product. All PCR products were resolved on a 1.5 % agarose gel (w/v).

117

and the DM-R / DM-PRENK product was probed with a DM-F / DM-PRENK PCR product, (Figure 5.4.B). The *Dmt* transgene was transcribed in all of the lines both in the flanking region and through the repeat. However, there were differences in the tissue specificity and transcription levels between the lines. *Dmt*-D showed high levels of transcription in all tissues, with bands visible on an agarose gel. Lines *Dmt*-B, -C and -E exhibited tissue specificity with transcription levels lower than that of the *Dmt*-D tissues. These bands were visible after Southern hybridisation. Transcription was observed in the brain in all lines, as well as the heart of *Dmt*-C.

5.2.2 Transcription in multiple tissues of *Dmt*-D

Transcription in *Dmt*-D appears to be transcribed in the three tissues studied. To investigate this further and any possible link between repeat stability levels, total RNA was made from a range of tissues from a 6 week old male mouse with 173 repeats. These included the quadriceps, total brain, liver, kidney, lung, heart, spleen, colon and testes (Figure 5.5.A). To investigate any age effects that may occur the quadriceps, total brain, liver, kidney, lung, heart, spleen, colon, testes and tongue from a mature *Dmt*-D mouse aged 14 months with 159 repeats were also examined (Figure 5.5.B). It appears that the *Dmt*-D transgene is ubiquitously transcribed in all of the examined tissues and that the level of transcription remains at a high level independent of the age of the mouse. It is likely that *Dmt*-D has integrated downstream of the promoter region of a housekeeping gene, although nothing is known about its proximity to the transgene.

5.2.3 Transcription in multiple tissues of *Dmt*-E

In order to determine whether any tissues other than the brain were transcribed multiple tissues from a mature *Dmt*-E male of 20 months with a repeat size of 162, were also examined. Although the quadriceps, total brain, spleen, liver, kidney, lung, heart, eye, colon, testes and tongue were all examined, the *Dmt*-E transgene was transcribed only in brain, lung, eye and testes (Figure 5.6).
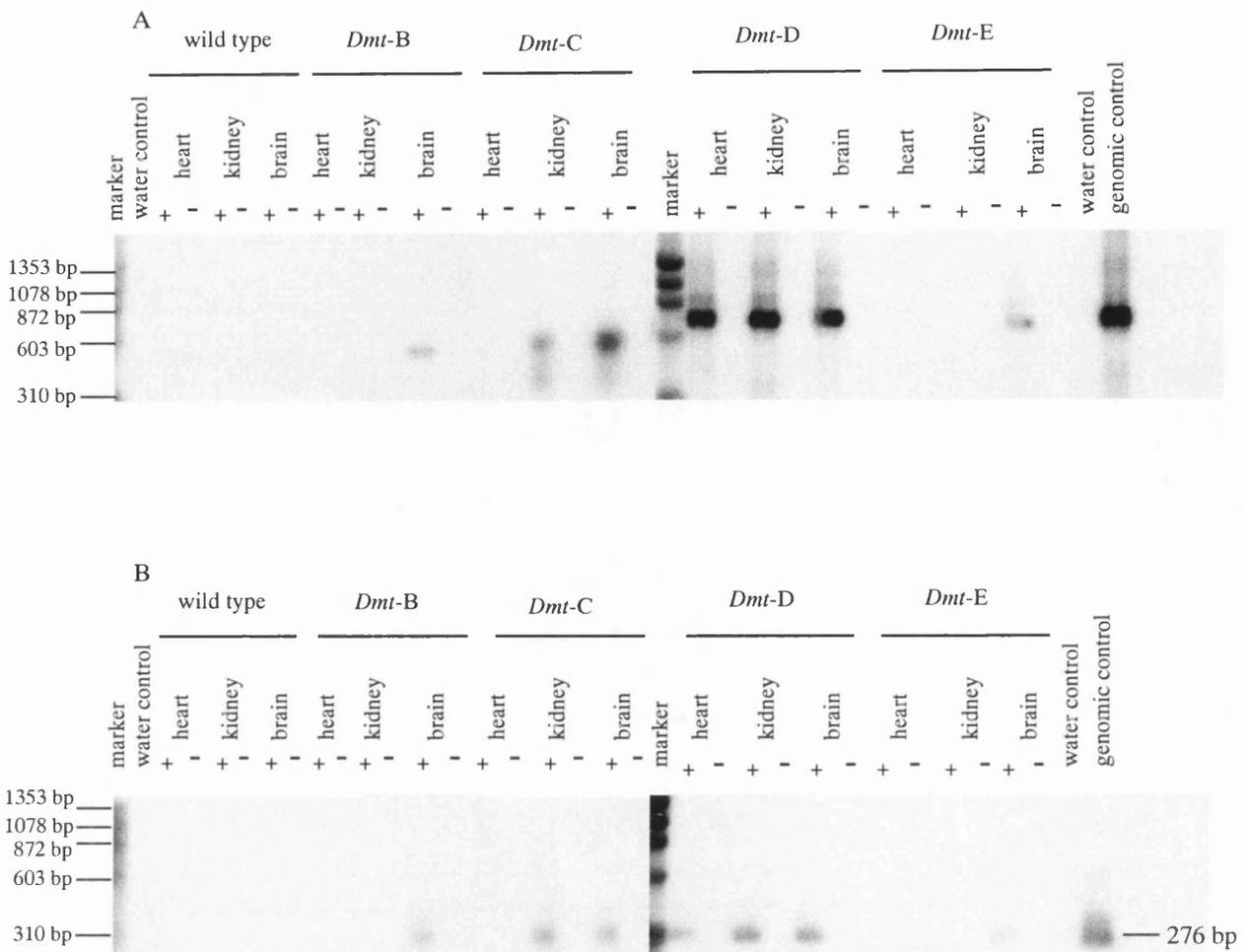
Figure 5.4
Transcription of transgene in *Dmt*-B, -C, -D and -E
To determine whether the transgene is transcribed in the *Dmt* mice, total RNA was made from the heart, kidney and brain of a 2 month old wild type FVB/N, a 20 month old *Dmt*-B mouse with 133 repeats, a 3 month old *Dmt*-C mouse with 159 and 162 repeats, a 4 month old *Dmt*-D mouse with 184 repeats and a 2 month old *Dmt*-E mouse with 158 repeats. 2.5 μg of total RNA from each tissue was used as a template for the generation of cDNA, which was primed using 50ng of random prime hexamers. The presence of the reverse transcriptase enzyme in the reaction is indicated by + (with), or - (without). 1μl of cDNA product was used as the template for RT-PCR using either A: DM-H / DM-DR to amplify across the transgenic repeat or B: DM-R / DM-PRENK which amplifies a non repetitive region of the transgene, 276 bp in size. All PCR products were resolved on a 1.5 % agarose gel before transfer to a nylon membrane and hybridisation with A: DM56 or B: DM-F / DM-PRENK PCR product.

119

A

6 week *Dmt*-D male
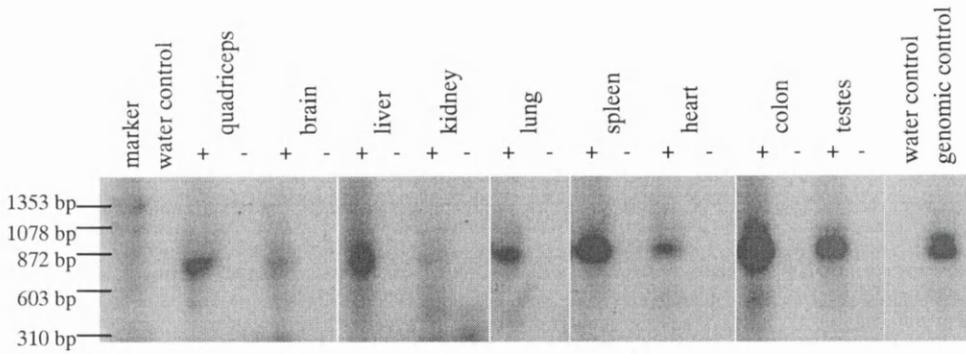
B

14 month *Dmt*-D male
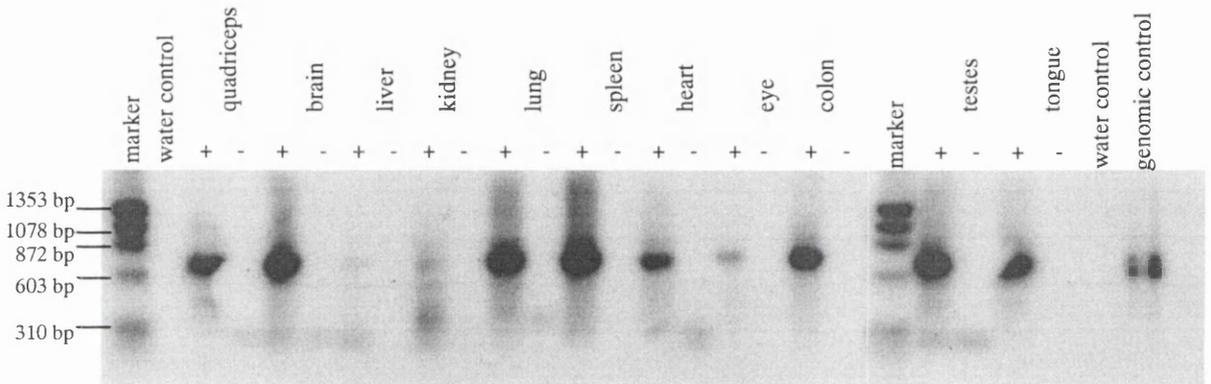
Figure 5.5
Transcription of transgene in multiple tissues of *Dmt*-D
To determine whether the transgene is transcribed in all tissues of *Dmt*-D and if the age of the mouse effects this transcription. RNA was prepared from the quadriceps, brain, liver, kidney, lung, spleen, heart, colon and testes of a 6 week old mouse with 179 repeats and the quadriceps, brain, liver, kidney, lung, spleen, heart, eye, colon, testes and tongue of a 14 month old mouse with 166 repeats.

2.5 µg of total RNA from each tissuewas used as a template for the generation of cDNA, which was primed using 50ng of random prime hexamers. The presence of the reverse transcriptase enzyme in thereaction is indicated by + ,(with) or -, (without). 1µl of cDNA product was used as the template for RT-PCR using either DM-H / DM-DR to amplify across the transgenic repeat. All PCR products were resolved on a 1.5 % agarose gel before transfer to a nylon membrane and hybridisation with DM56.
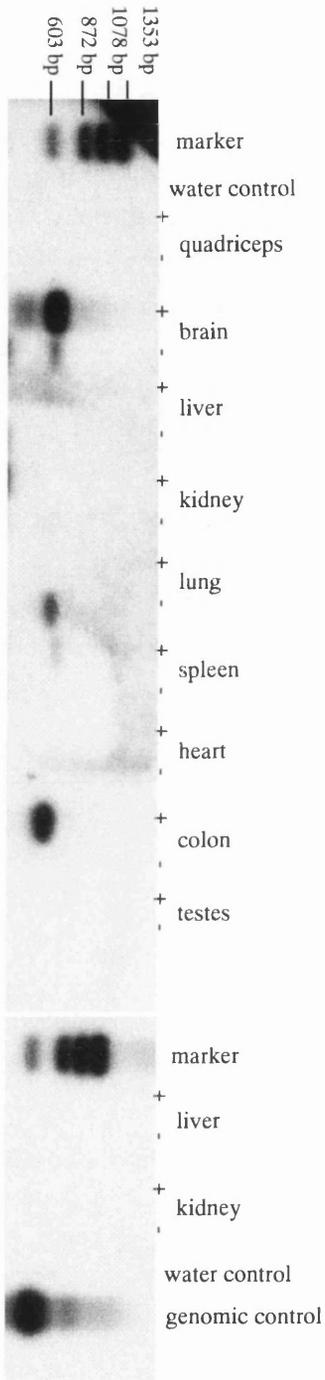
Figure 5.6
Transcription of transgene in multiple tissues of *Dmt*-E
To determine which tissues of *Dmt*-E the transgene is transcribed in a selection of tissues were examined from a 20 month old male mouse with a repeat size of 162. RNA was prepared from the quadriceps, brain, liver, kidney, lung, spleen, heart, eye, colon, testes and diaphragm. 2.5 µg of total RNA from each tissuewas used as a template for the generation of cDNA, which was primed using 50ng of random prime hexamers. The presence of the reverse transcriptase enzyme in the reaction is indicated by + ,(with) or - ,(without). 1µl of cDNA product was used as the template for RT-PCR using DM-H / DM-DR to amplify across the transgenic repeat. All PCR products were resolved on a 1.5 % agarose gel before transfer to a nylon membrane and hybridisation with DM56.

5.3 Orientation of the *Dmt* transgenes within their integration site

To allow investigation of the orientation of the CTG repeats within their integration sites cDNA was generated using 2.5 µg of total brain RNA from each line, *Dmt*-B, -C, -D, and -E. The cDNA was primed with 0.045 mM of GAPDHA and either 0.045 mM of DM-H or 0.045 mM of DM-DR in the final reaction volume. The cDNA was amplified with GAPDHA / GAPDHBR to check the success of the reverse transcriptase reaction (data not shown) and with DM-H / DM-DR to determine the orientation of the repeat from this pool of first strand cDNA. In an attempt to corroborate the preceding experiment cDNA was also generated using 0.045mM of GAPDHA and either 0.045mM of DM-R or 0.045mM of DM-PRENK in the final reaction volume. The ability to amplify the PCR product in one direction and not the other would indicate the orientation that the transgene had integrated in. The DM-H / DM-DR amplification was inconclusive as multiple bands were present on the 1.5% agarose gel (w/v) (Figure 5. 7.A) with patterns apparently common to the primer the cDNA was generated with. There were 2 distinct band patterns depending on orientation. Each band was the same size in each mouse. These visible bands did not hybridise with DM56. Other bands did however, hybridise with DM56 (Figure 5.7.B). These bands were probably transgenic in origin, an assumption made because of the apparent size difference present in the different mice. They also occurred in both orientations. DM-H is a 20mer and DM-DR is a 20mer with annealing temperatures of 74 °C and 68 °C respectively. It is likely that these primers are not specific enough and that the PCR product has been primed at an alternative site during the reverse transcriptase reaction which was incubated at the recommended temperature of 42 °C.

The DM-R / DM-PRENK amplification when hybridised with the DM-F / DM-PRENK PCR product also revealed bands (Figure 5.8.B) not visible on the 1.5% agarose gel (w/v) (Figure 5.8.A). These single bands when present were present in the CAG orientation in *Dmt*-C, -D and -E and the CTG orientation in *Dmt*-E and –D. However, although a bias of band intensity was visible in the in the CAG orientation for *Dmt*-D, it was difficult to draw conclusions from this experiment.

It was suggested that the reverse transcriptase enzyme could function at 55 °C, but attempts to generate transgene specific cDNA at this temperature were unsuccessful (data not shown). As an alternative, transgenic 12mer primers were designed with annealing
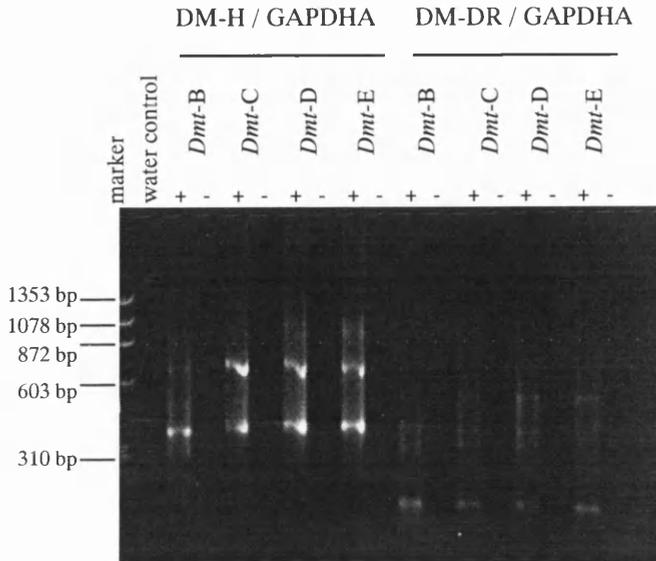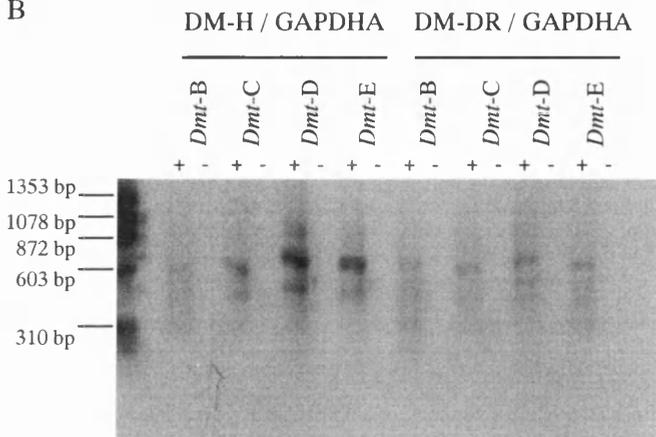
Figure 5.7
First strand DNA generated with DM-H or DM-DR is used to determine the orientation of integration of the *Dmt* transgenes

2.5 μg of total RNA from the brain of a 2 month old wild type FVB/N, a 20 month old *Dmt*-B mouse with 133 repeats, a 3 month old *Dmt*-C mouse with 159 and 162 repeats, a 4 month old *Dmt*-D mouse with 184 repeats and a 2 month old *Dmt*-E mouse with 158 repeats was used as a template for the generation of cDNA.

First strand cDNA was generated for each tissue using either 0.045 mM of DM-H and 0.045 mM of GAPDHA or 0.045 mM of DM-DR and 0.045 mM of GAPDHA in the final reaction volume. The presence of the reverse transcriptase enzyme in the reaction is indicated by + (with), or - (without).

1μl of each cDNA product was used as the template for RT-PCR amplification with DM-H / DM-DR the PCR products were resolved on (A) a 1.5 % agarose gel (w/v) before transfer to a nylon membrane and (B) hybridisation with DM56.

123
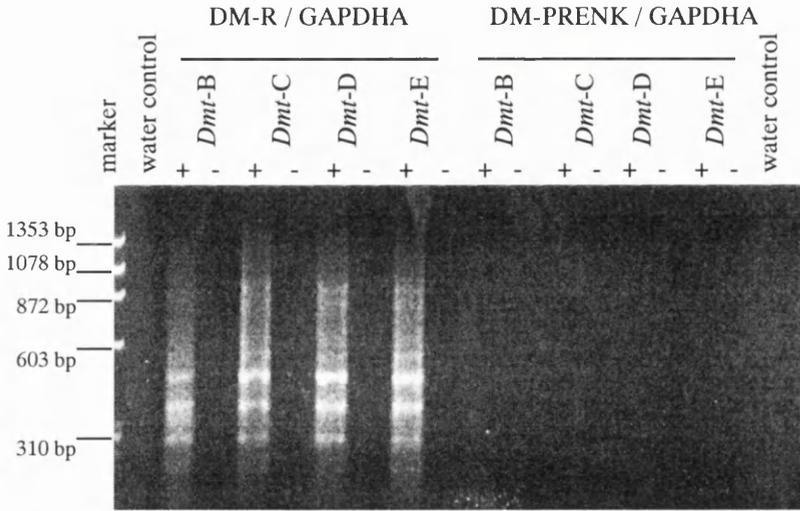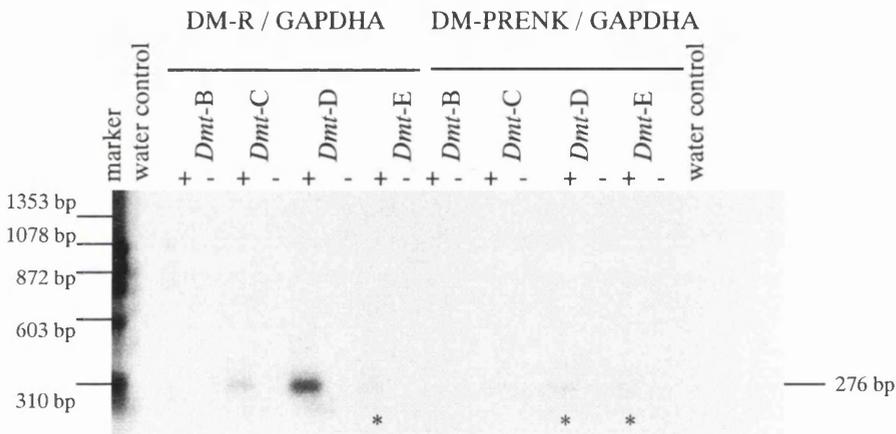
Figure 5.8
First strand DNA generated with DM-R or DM-PRENK are used to determine the orientation of integration of the *Dmt* transgenes
2.5 μg of total RNA from the brain of a 2 month old wild type FVB/N, a 20 month old *Dmt*-B mouse with 133 repeats, a 3 month old *Dmt*-C mouse with 159 and 162 repeats, a 4 month old *Dmt*-D mouse with 184 repeats and a 2 month old *Dmt*-E mouse with 158 repeats was used as a template for the generation of cDNA.
First strand cDNA was generated for each tissue using either 0.045 mM of DM-R and 0.045 mM of GAPDHA or 0.045 mM of DM-PRENK and 0.045 mM of GAPDHA in the final reaction volume. The presence of the reverse transcriptase enzyme in the reaction is indicated by + (with), or - (without).
1μl of each cDNA product was used as the template for RT-PCR amplification with DM-R / DM-PRENK, the PCR products were resolved on (A) a 1.5 % agarose gel (w/v) before transfer to a nylon membrane and (B) hybridisation with DM-F / DM-PRENK PCR product.

124

temperatures of 40 °C. It was hoped that DM-H12 and DM-DR12 would improve the specificity of the reverse transcriptase reaction. Transgene specific cDNA was generated using 2.5 μg of total RNA from the brain of a wild type FVB/n and *Dmt*–B, -C, -D, and –E mice, with 0.045mM GAPDHA and either 0.045mM DM-H12 or 0.045mM DM-DR12 in the final reaction volume. RT-PCR was carried out with both GAPDHA / GAPDHBR and DM-H / DM-DR.

Again the agarose gel (Figure 5.9 A) was uninformative, but transfer of the DM-H / DM-DR PCR products to nylon membrane and probing with DM56 revealed single bands of transgenic size for each line. Although bands could be observed in both orientations in all four examined lines, there was a bias in intensity in each case. It is thought that *Dmt*-C, -D, and -E are transcribed primarily in the CAG orientation and that *Dmt*-B is primarily transcribed as a CTG repeat (Figure 5.9 B).

It was suggested that the GAPDHA primer could be priming across the transgenic repeat in both directions and this would account for the presence of bands in both orientations. To further investigate the effect of the GAPDHA, 3 further cDNAs were generated using 2.5 μg of *Dmt*-D brain total RNA, with 0.045mM DM-H12 or 0.045mM DM-DR12 or 0.045mM GAPDHA in the final reaction volume of 11.5 μl. RT- PCR was carried out on each cDNA using DM-H / DM-DR. Each PCR product was resolved on a 1.5% agarose gel (w/v) before transfer to a nylon membrane. Southern hybridisation of the cDNAs with DM-56 revealed a strong band with the DM-H12 cDNA product and a weaker band for the DM-DR12 cDNA RT-PCR product. The band which hybridised with the GAPDHA cDNA RT-PCR product was weaker still (Figure 5.10). This suggests that the presence of GAPDHA in the RT-PCR reaction when generating the orientation specific first strand cDNA was contributing to the band which hybridised with DM56. However, this experiment would appear to confirm that although *Dmt*-D is transcribed in both orientations it is primarily as a CAG transcript.

5.4 Investigation of the presence of the transcribed transgenes in the cytoplasm

Experiments to ascertain whether the transcribed transgenes were present in the cytoplasm or restricted to the nucleus of the 4 transcribed lines were carried out. If the transgenes were

125
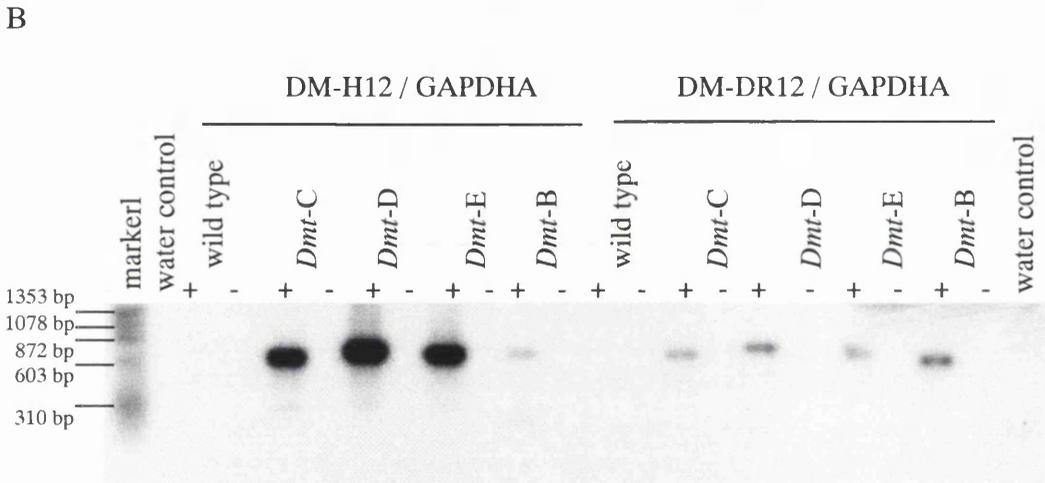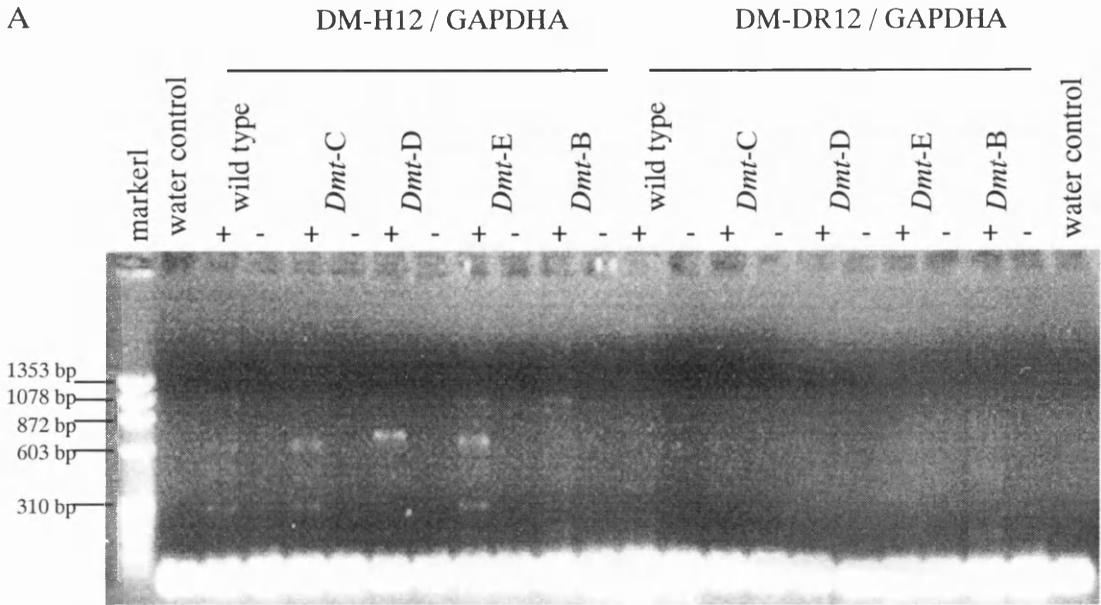
Figure 5.9
Determination of the orientation of the *Dmt*-transgene by the ability to amplify the transgene from pools of first strand DNA generated with DM-H12 or DM-DR12
2.5 µg of total RNA from the brain of a 2 month old wild type FVB/N, a 20 month old *Dmt*-B mouse with 133 repeats, a 3 month old *Dmt*-C mouse with 159 and 162 repeats, a 4 month old *Dmt*-D mouse with 184 repeats and a 2 month old *Dmt*-E mouse with 158 repeats was used as a template for the generation of cDNA.
First strand cDNA was generated for each tissue using either 0.045 mM of DM-H12 and 0.045 mM of GAPDHA or 0.045 mM of DM-DR12 and 0.045 mM of GAPDHA in the final reaction volume. The presence of the reverse transcriptase enzyme in the reaction is indicated by + ,(with) or -, (without).
1µl of each cDNA product was used as the template for RT-PCR amplification with DM-H / DM-DR and the PCR products were resolved on (A) a 1.5 % agarose gel (w/v) before transfer to a nylon membrane and (B) hybridisation with DM56.

126
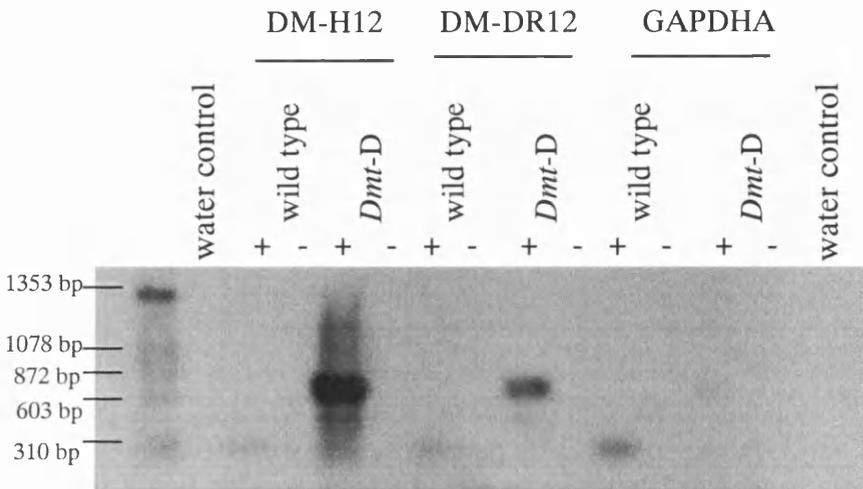
Figure 5.10
To determine whether the primer GAPDHA could generate a transgenic PCR product
2.5 µg of total RNA from the brain of a 4 month old *Dmt*-D mouse with 184 repeats was used as a template for the generation of cDNA. First strand cDNA was generated for each tissue using either 0.045 mM of DM-H12 or 0.045 mM of DM-DR12 or 0.045 mM of GAPDHA in the final reaction volume. The presence of the reverse transcriptase enzyme in the reaction is indicated by + (with), or - (without). 1µl of each cDNA product was used as the template for RT-PCR amplification with DM-H / DM-DR,the PCR products were resolved on a 1.5 % agarose gel (w/v)(not shown) before transfer to a nylon membrane and hybridisation with DM56 .

127

found only in the nuclear fractions it could be assumed that they were located in intronic regions of the genome or that the integration of the repeat may have disrupted transcript processing in some way. The repeat may have made the export of the transcript out of the nucleus sterically unviable and it had therefore become trapped (Koch and Leffert, 1998).

5.4.1 The enrichment of polyA RNA from total RNA by using Dynabeads™ Oligo (dT)$_{25}$

100 µg of total RNA from wild type, *Dmt*-B, -C, -D and -E mouse brains were enriched for polyA RNA by passing through Dynabeads™ Oligo (dT)$_{25}$ according to the Dynabeads™ protocol. 1 µl of this polyA enriched RNA was used as template for cDNA generated with 0.09 mM of random prime hexamers in the final reaction volume. This cDNA was used as a template for RT-PCR carried out with GAPDHA / GAPDHBR which confirmed the success of the reverse transcriptase. Amplification across the repeat with DM-H / DM-DR resulted in a transgenic band in all of the transgenic lines. However, amplification with CLW2F / CLW4R revealed the presence of a heteronucleic RNA sized product as well as a spliced products (Figure 5.11). It is possible that this single round of enrichment was not enough to exclude none polyA products. Alternatively successful enrichment may have been achieved but the pool of polyA products included polyadenylated transcripts, where not all of the introns had been removed by splicing as well as messenger RNA. Transcription levels in all lines except *Dmt*-D require Southern hybridisation to be detectable and this sensitive method could be detecting heteronucleic product that has already acquired a polyA tail. To continue this study alternative strategies were adopted.
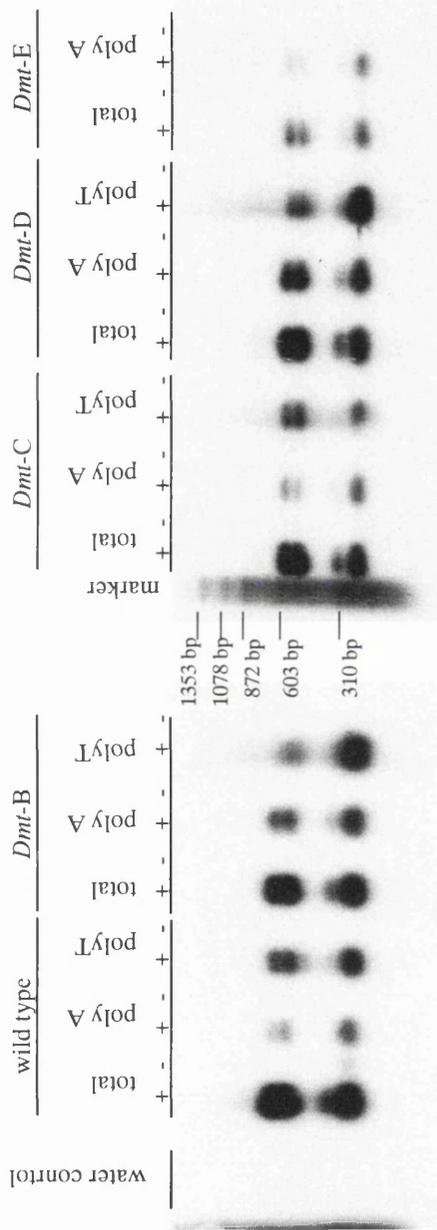
Figure 5.11

Investigation of the transcribed transgene in the cytoplasm

2.5 μg of total RNA from the brain of a 2 month old wild type FVB/N, a 20 month old *Dmt*-B mouse with 133 repeats, a 3 month old *Dmt*-C mouse with 159 and 162 repeats, a 4 month old *Dmt*-D mouse with 184 repeats and a 2 month old *Dmt*-E mouse with 158 repeats was used as a template for the generation of cDNA. First strand cDNA was generated for each tissue using either 50 ng of random prime hexamers (total), or 0.09mM of the XYT primer (poly T) Friedrich *et al.* (1993) in the final 11.5 μl reaction volume.

2.5 μg of total RNA from the brain of a 2 month old wild type FVB/N, a 20 month old Dmt-B mouse with 133 repeats, a 3 month old Dmt-C mouse with 159 and 162 repeats, a 4 month old Dmt-D mouse with 184 repeats and a 2 month old Dmt-E mouse with 158 repeats was enriched for mRNA by using Dynabeads Oligo (dT)$_{25}$. First Strand cDNA synthesis was carried out using on 1μl of the polyA enriched RNA fraction (polyA).

The presence of the reverse transcriptase enzyme in the reaction is indicated by + ,(with) or -, (without).

1μl of each cDNA product was used as the template for RT-PCR amplification with CLW2F / CLW4R.The PCR products were resolved on a 1.5 % agarose gel (w/v) before transfer to a nylon membrane and hybridisation with CLW2F / CLW4R genomic probe. The presence of multiple heteronucleic sized products as wells as the mRNA product was revealed.

5.4.2 First Strand cDNA synthesis with XYT primer

First Strand cDNA synthesis was carried out on 2.5 μg of total RNA isolated from the brain of wild type, *Dmt*-B, -C, and -D mice using 0.09mM of the XYT primer (Friedrich *et al.*, 1993) in the final 11.5 μl reaction volume. The XYT primer is a polyT primer designed to anneal to RNA transcripts with a polyA tail. The success of the reverse transcriptase was confirmed by amplifying GAPDHA / GAPDHBR. The amplification of the repeat resulted in a band in the transgenic lines. Again however, amplification between introns 2 and 4 revealed the presence of multiple products as well as the product of the mRNA template (Figure 5.11).

To determine whether transcripts containing the transgene are exported into the cytoplasm a method to separate the nuclear and cytoplasmic RNAs would be necessary.

5.4.3 Separation of cytoplasmic and nucleic cell fractions

To separate nuclear and cytoplasmic fractions with complete confidence of having eliminated cross contamination was thought to be difficult when using tissues. Separation of the nuclear fraction by centrifugation could be achieved, but the cytoplasmic fraction might still be contaminated by nuclei lysed in the initial tissue homogenisation and only complete confidence in the purity of a pure cytoplasmic fraction could conclusively demonstrate that the transgenes were exported to the cytoplasm.

Alternatively cell lines can be lysed using a gentle chemically based method that allows the purification of cytoplasmic and nuclear RNA fractions. Cell lines were available from *Dmt*-D mice only. The chosen cell line which was established from the kidney of a 5 week *Dmt*-D male had undergone few passages and this limited the possibility of rearrangements. A flask containing 3×10^6 cells from the *Dmt*-D 3111 kidney cell line were separated by a modified method (Hamshere *et al.*, 1997) and RNA extracted using TRI REAGENT™.

0.09 mM of random prime hexamers in the final reaction volume was used to make cDNA using either 2.0 μg of cytoplasmic or 2.0 μg nucleic RNA as template. Reverse transcriptase PCR was carried out with the control primers GAPDHA / GAPDHBR. Amplification across exon 2 and 4 of DMPK revealed both a heteronucleic and spliced product from the nuclear fraction. However, only a spliced product was present when the

cytoplasmic fraction was hybridised. The amplification across the CTG repeat using both cDNAs as template resulted in a band from the nuclear fraction only (Figure 5.12).

5.5 Conclusion

To summarise the results of this chapter all 4 of the *Dmt* transgenic lines examined are transcribed, although there is a difference in the nature of the transcription of these lines. Although the level of transcription has not been determined exactly, equivalent amounts of RNA from each line was used to generate the cDNA for each set of experiments. Equal volumes of the generated cDNA from each line and tissue were used as the template for each of the RT-PCR amplifications.

Dmt-B, -C and -E exhibit transcription in certain tissues with variation between the tissue types, that is transcription is detected in some tissues and not at all in others. *Dmt*-D however, demonstrates a higher level of transcription than that detected in the examined tissues of the other lines and the transcription is apparently ubiquitous. The transcription level observed in the kidney appears no different to that of the other examined tissues.

Investigation of the orientation of the transcribed repeats within their integration sites suggests *Dmt*-C, -D and -E are transcribed as CAG repeats, while *Dmt*-B has integrated in the CTG orientation. This eliminates the suggestion differences in the repeat orientation between *Dmt*-D and the other lines could account for the gross somatic instability associated with *Dmt*-D.

Attempts to determine whether the repeats are present in the cytoplasmic RNA fraction have been inconclusive for lines *Dmt*-B, -C and -E. However, it has been determined that the transcribed repeat is not present in the mRNA of the *Dmt*-D kidney cell line. This suggests that the *Dmt* transgene has probably integrated into an intron. Alternatively the integration site could be within a UTR but the expanded repeat could be preventing the export of the transcript into the cytoplasm (Koch and Leffert, 1998). The possibility of the integration site being within an exon is less likely. Again the repeats presence may sterically inhibit export. However, such an integration site would also effect the protein into which the transgene has integrated and no resulting phenotype has been observed in the *Dmt*-D mice. *Dmt*-D mice have survived as homozygotes. The arguments for UTR and exonic integration sites are again
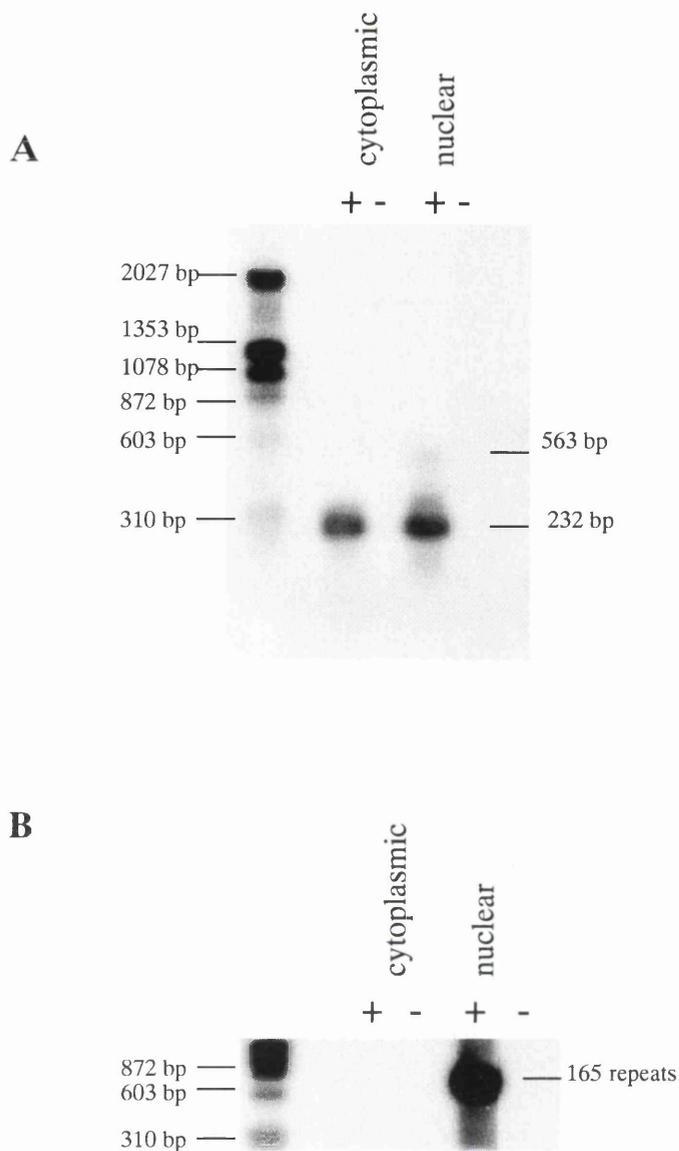
Figure 5.12
To determine the presence of the Dmt-D transgene in cytoplasmic and nuclear RNA fractions

Cytoplasmic and nuclear cDNA was generated with 2µg of each template RNA isolated from the Dmt-D 3111 kidney cell line, with a 165 repeat size, using 50 ng of random prime hexamers. The presence of the reverse transcriptase enzyme in the reaction is indicated by + ,(with) or -, (without). 1µl of each cDNA product was used as the template for RT-PCR A. Amplification of the cDNA with Dmpk primer pair hybridised with Dmpk PCR product and B. Amplification of the cDNA across the transgenic repeat hybridised with DM56. The PCR products were resolved on a 1.5 % agarose gel (w/v) before transfer to a nylon membrane and hybridisation with (A) Dmpk PCR product and (B) DM56 .

weakened with the observation that the expansion within the *Dmt* transgene may not be so large that their export into the cytoplasm would be prevented.

It has been hypothesised that part of the multifaceted nature of the instability exhibited by DM1 patients could be attributed to the transcribed CUG DM1 repeat forming hairpin structures. Such structures are retained within the nucleus and might act as protein "sinks" removing certain proteins from their correct cellular function (Phillips *et al.*, 1998; Tian *et al.*, 2000). The grossly expanded *Dmt*-D transcribed repeats are CAG tracts and would not cause proteins that may be associated with DM1 phenotype to aggregate.

In conclusion although all of the *Dmt* lines are transcribed the major difference revealed between the somatically unstable line is in the nature of the transcription. It is possible that differences in the integration sites, i.e. the conformation of the DNA or transcription may account for the differences in stability observed in the somatic tissues between *Dmt*-D and the other lines. However, the significance of such an observation may be limited as the unstable repeat tract in *DMPK* is transcribed in a tissue specific manner.

Although the transcription of the repeat associated with DM1 is associated with the disease phenotype (Chapter 1.8), transcription does not appear to play a major role in repeat instability. None of the *Dmt* transgenic lines display a DM1 associated phenotype.

Chapter 6

Investigation of the stability levels of the *Dmt* transgene on a background deficient in a mismatch repair protein

6.1 Introduction

The experiments in the previous chapters have shown that the *Dmt*-D transgenic repeat exhibits tissue specific expansion biased instability and somatic mosaicism. The instability does not appear to be related to the level of transgenic transcription or indeed the cell turnover of a particular tissue. Hypothesised mechanisms of instability involve the aberrant behaviour of a functioning mismatch repair (MMR) protein complex mistakenly interpreting misaligned repeat regions (Figure 3.1).

HNPCC kindreds are associated with the mutation in one of the genes associated with the MMR protein complex. Studies into the stability of microsatellite DNA in HNPCC cases showing a germline mutation in a mismatch repair gene also showed microsatellite instability in their tumours (Liu, 1996). While a study of the trinucleotide repeat of SBMA, SCA type 1 and DM1 in early breast carcinoma tumours found triplet repeat instability in around 5% of the tumours. If instability was detected, it was present at all three loci, and among these three it was the DM1 locus which exhibited the greatest degree of instability (Shaw *et al.*, 1996), demonstrating that although the MMR protein complex has a role to play in repeat instability, the level of instability shown is effected by *cis* acting factors. Accepting that the MMR protein complex does play a role in repeat instability, then what is the effect of a damaged MMR complex on an expanded triplet repeat, for example the *Dmt* transgene?

The MMR protein complex, role involves 'proofreading' DNA, this includes correcting errors in base pairing, unpaired regions and the removal of unusual DNA structures, which includes hairpins or loopout structures (Umar *et al.*, 1994; Yu *et al.*, 1995). *MutLα* is a heterodimer that forms part of the MMR protein complex. It consists of *MLH1* and *PMS2* and mutations in *PMS2*, account for a small number of HNPCC cases (Buermeyer *et al.*, 1999). Baker *et al.* (1995) knocked out the function of *Pms2* in mice by targeting the insertion of a neomycin cassette into the *Pms2* locus (Figure 6.1). Initial observation of these *Pms2* (-/-) mice showed the males to be sterile. Problems were observed in chromosome alignment and pairing indicating a role in meiosis (Baker *et al.*, 1995), although sterility was also observed in male

1. Amplification of PCR product from wildtype *Pms2*.

2. Insertion of the neomycin cassette into *Pms2* knocks out the gene function and disrupts the mP2-3 primer site.

3. Amplification of PCR product from *Pms2* knockout.
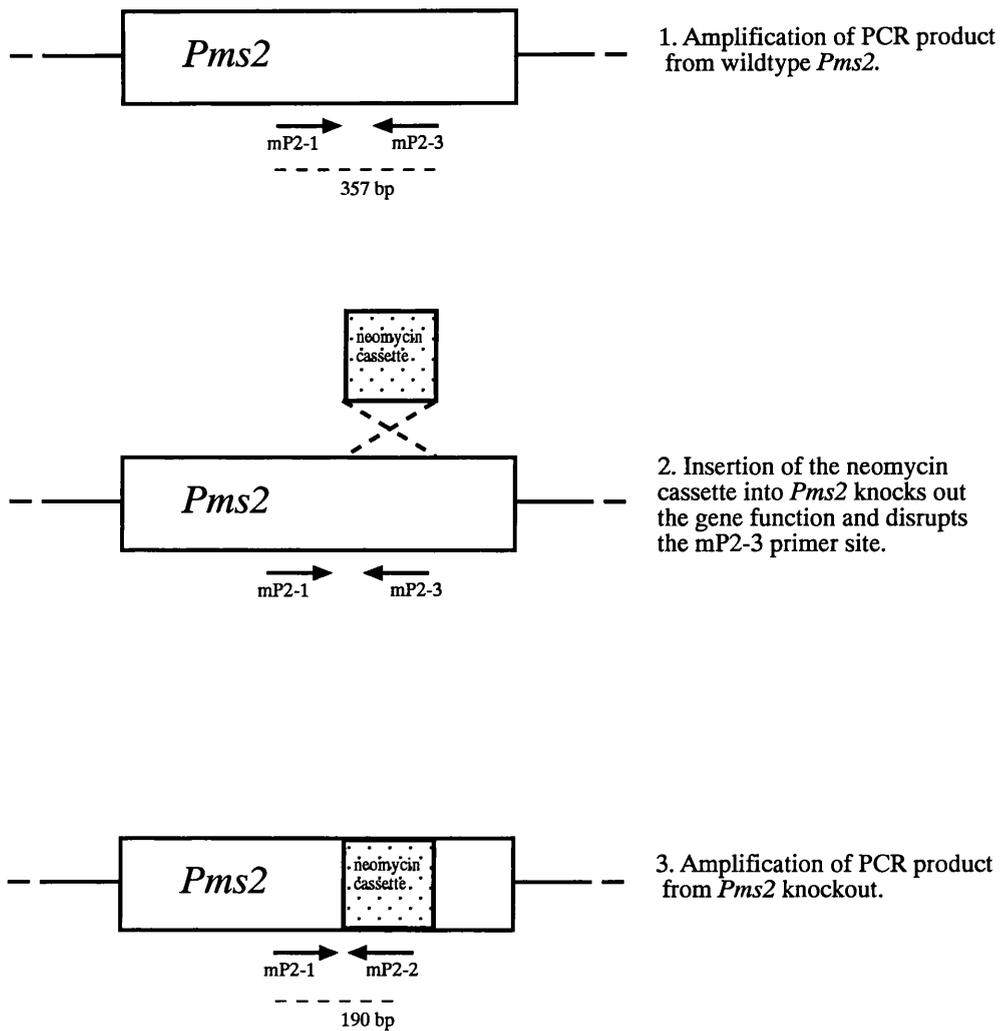
Figure 6.1
The *Pms2* knockout
The targeted insertion of a neomycin cassette into the *Pms2* gene knocks out the function of the protein.

and female *Mlh1* (-/-) mice (Baker *et al.*, 1996). The *Pms2* (-/-) female mice were observed to be fertile as were male and female *Pms2* (+/-) mice. A study of *Pms2* (-/-) mice observed 95% of the 20 mice studied developed tumours by 6 months, the majority of which were lymphomas and the remainder sarcomas (Qin *et al.*, 1999). Comparison of the rate of mutation in long and short tracts of mononucleotide in *Mlh1* (-/-) or *Pms2* (-/-) murine null homozygotes found the mutation rate in *Mlh1* (-/-) mice to be 2-3 times higher than in *Pms2* (-/-) mice (Yao *et al.*, 1999). To study the effect of the MMR complex on *Dmt* repeat stability, *Dmt*-D and *Dmt*-E mice were crossed with mice deficient in *Pms2*.

6.2 Mouse typing for the *Pms2* status of each mouse

To determine the *Pms2* status of each mouse, 1 μl (~1 ng) of tail dilution was set up in a 7 μl PCR reaction containing 1 X PCR buffer, 0.35 units of *Taq* polymerase and 1 μM each of mP2-1, mP2-2 and mP2-3 and cycled 30 times under standard conditions, with an annealing temperature of 61 °C. The reaction volume was made up to 10 μl with loading dye and the complete volume resolved on a 1 % agarose, 2 % nusieve gel (w/v) at 100 V for ~ 2 hours before photographs were taken. The mP2-1 and mP2-3 primer combination will amplify a band 357 bp in size from mice containing an undisrupted copy of *Pms2*. The mP2-1 and mP2-2 primer combination will amplify a band of 190 bp in mice whose copy of *Pms2* has been disrupted by the insertion of a neomycin cassette into exon 2. This disruption results in a null mutation (Baker *et al.*, 1995). This mP2-1, mP2-2 and mP2-3 primer mix allows the identification *Pms2* homozygotes (+/+), *Pms2* heterozygotes (+/-) and *Pms2* null homozygotes (-/-) (Figure 6.2). Although the PCR amplifications result in extra bands of a similar size to those of interest, the identity of the correct bands was confirmed by the sequencing of the PCR products resulting from the primer mix combination. After resolving the PCR products on a 2% agarose gel, each band of interest was isolated, the products were purified from the agarose gel using the Qiagen PCR Purification Kit. Each isolated product was cloned into the pCR®2.1-TOPO vector using the TOPO TA Cloning kit™ as described in the kit's protocol and transformed into the F*Escherichia coli* TOP 10 bacterial strain. Resulting white colonies were grown up overnight in LB and the plasmid isolated using the Qiagen Plasmid DNA Mini Prep Kit. The PCR products were then sequenced using plasmid specific primers and the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit.
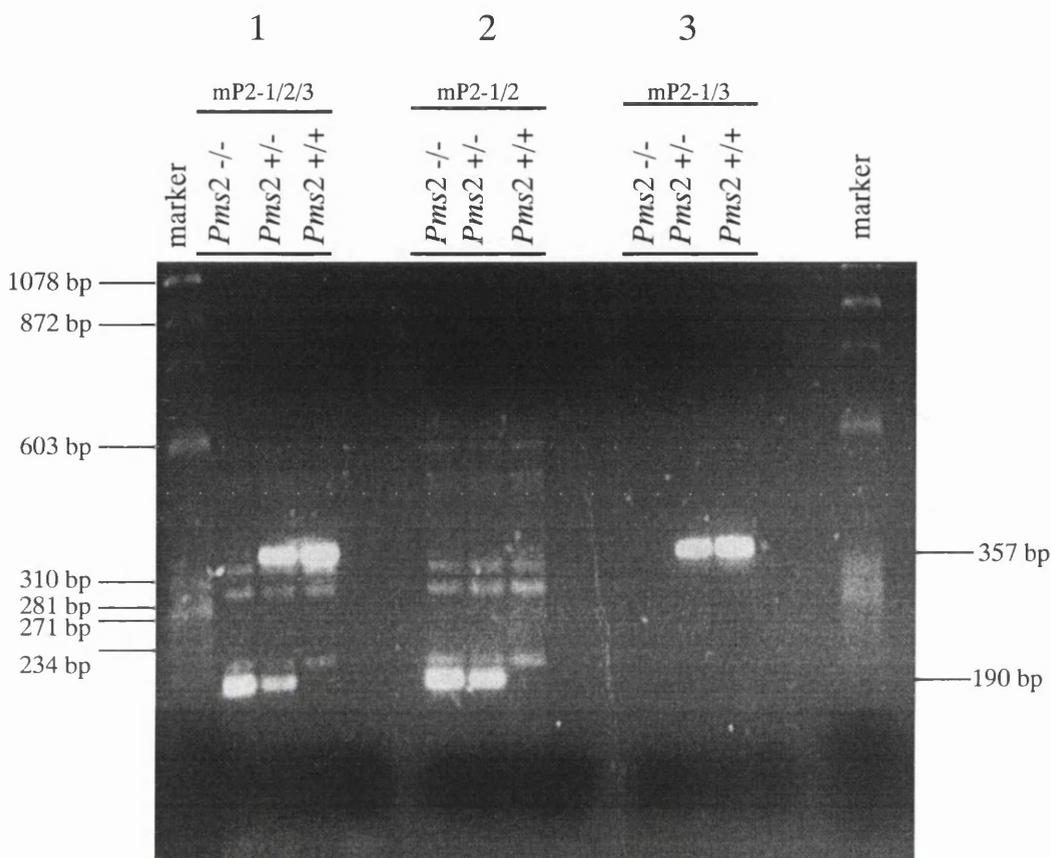
Figure 6.2
Identification of *Pms2* transgenic mice
PCR amplification to detect *Pms2* homozygotes (+/+), hemizygotes (+/-) or null homozygotes (-/-) by amplifying 1μl of tail lysate (~ 1 ng of DNA) in a 7μl volume containing1 μM each of (1) mP2-1, mP2-2 and mP2-3, or (2) mP2-1, mP2-2 or (3) mP2-1, mP2-3. These primer combinations will amplify (1) both the positvie, 357bp and the negative, 190 bp, *Pms2* products. (2) The 190bp *Pms2* product. (3) The 357bp *Pms2* product. The PCR products were resolved on a 1 % agarose, 2 % nusieve gel (w/v) at 100 V for ~ 2 hours.

The bands identified in Figure 6.2 as the correct *Pms2* products were the only sequenced products that contained the correct primer combinations.

6.3 Matings between *Dmt* and *Pms2* mice

Matings were set up in the first instance between *Dmt*-D or *Dmt*-E mice hemizygous, (+/-) for the *Dmt* transgene and mice hemizygous (+/-) for the *Pms2* transgene. Matings were then set up from subsequent generations between *Dmt* (+/-) / *Pms2* (+/-) mice and mice transgenic for only *Pms2* (+/-). Although the *Pms2* status of the mice varied, (+/+), (+/-) and (-/-), all of the *Dmt* mice bred were hemizygous for the transgene. The number of *Dmt*-D (+/-) / *Pms2* (-/-) offspring generated was low, 20 mice, while only 2 *Dmt*-E (+/-) / *Pms2* (-/-) were generated. The study of *Dmt* (+/-), *Pms2* (-/-) was difficult as the majority became ill before 6 months of age, developing paralysis of the hind legs, a condition which lead to rapid deterioration and subsequent sacrifice of the mouse. Analysis of a *Dmt*-D (+/-), *Pms2* (-/-) mouse at the onset of this paralysis at 26 weeks and its apparently healthy *Dmt*-D (+/-), *Pms2* (+/-) littermate after their sacrifice at the same timepoint found no tumours present in either mouse. Both mice were found to have leukaemia. These investigations were inconclusive.

6.4 Somatic mosaicism in a *Dmt*-E transgenic mouse on a *Pms2* (-/-) background

The effect of the loss of the *Pms2* protein on the transgenic repeat in a stable line was examined by crossing the stable *Dmt*-E line onto a *Pms2* deficient background. A *Dmt*-E (+/-), *Pms2* (-/-) mouse, with a repeat size of 160 was sacrificed at 12 months because it had developed multiple muscle sarcomas. A selection of tissues were removed including the tumours and the associated muscle, as well as blood, heart, lung, kidney, colon and skin. DNA was prepared from the proteinase K lysate by phenol / chloroform extraction and the DNA concentration was determined by spectrophotometry before diluting to 1200 pgμl^{-1}, 120 pgμl^{-1} and 12 pgμl^{-1}. Duplicate SP-PCR amplifications were made using the 1200 pgμl^{-1} and the 12 pgμl^{-1} concentrations as templates. Instability of the transgenic repeat was observed in some of the tissues, including but not exclusively the tumours. The observed instability was deletion

biased (Figure 6.3). This degree of instability had only been observed previously in the tumour associated tissues and blood of a 25 month *Dmt*-E mouse (Chapter 3.4.4).

6.5 Somatic mosaicism in *Dmt*-D, transgenic mice on a *Pms2* (-/-) or *Pms2* (-/+) background

A male *Dmt*-D (+/-), *Pms2* -/- mouse, with a transgenic repeat size of 163 was sacrificed at 6 months, after it had developed a tumour next to the spinal cord. Its female *Dmt*-D, *Pms2* +/-, littermate, with a transgenic repeat size of 188 was sacrificed at the same timepoint. DNA was prepared from a selection of tissues using proteinase K lysis. The tissue lysates were diluted assuming that 1 μl of the lysate contained ~ 1 ng of DNA. Serial dilutions 1/10 (~ 1000 pgμl⁻¹), 1/100 (~ 100 pgμl⁻¹), 1/1000 (~ 10 pgμl⁻¹) were carried out in dilution buffer containing 0.1μM DM-H. 1μl of the 1/10 and 1/1000 serial dilutions were used as templates for duplicate and quadruplicate small pool PCR amplifications respectively. All SP-PCR conditions were consistent with those in Chapter 3. Initial comparison between the tissues of the two mice reveals apparent differences between the kidney samples. The kidney that has been reported as the most unstable of the *Dmt*-D tissues (Chapter 3), with instability visible at 6 months. However, there is apparently a difference in stability in the *Dmt*-D, *Pms2* -/-, mouse when compared to its *Dmt*-D *Pms2* +/- littermate. The kidney in the *Dmt*-D, *Pms2* -/- mouse is apparently more stable. This means that any qualitative observations at this timepoint are very preliminary. At 6 months instability is only just becoming visible in tissues other than kidney as differences in tissue instability profiles begin to develop (Figure 6.4). Comparison between the heart and eye tissue of the littermates also suggests there may be differences in stability levels. However, no absolute conclusions could be drawn from these initial experiments, as quantification of the levels of stability of the different tissues was not carried out. At the time of this experiment these were the oldest mice available, and time restriction prevented quantification of the instability observed in the mice. Studies by Mr Mario Pereira and Miss Laura Ingram, (personal communication), carried out on 8 month old *Dmt*-D *Pms2* -/- and *Dmt*-D *Pms2* +/- littermates, which include quantification, observed that although the repeat does indeed appear to be more stable in the *Dmt*-D *Pms2* -/- littermate, with fewer expansion events. However, large deletion events appear to occur more frequently.
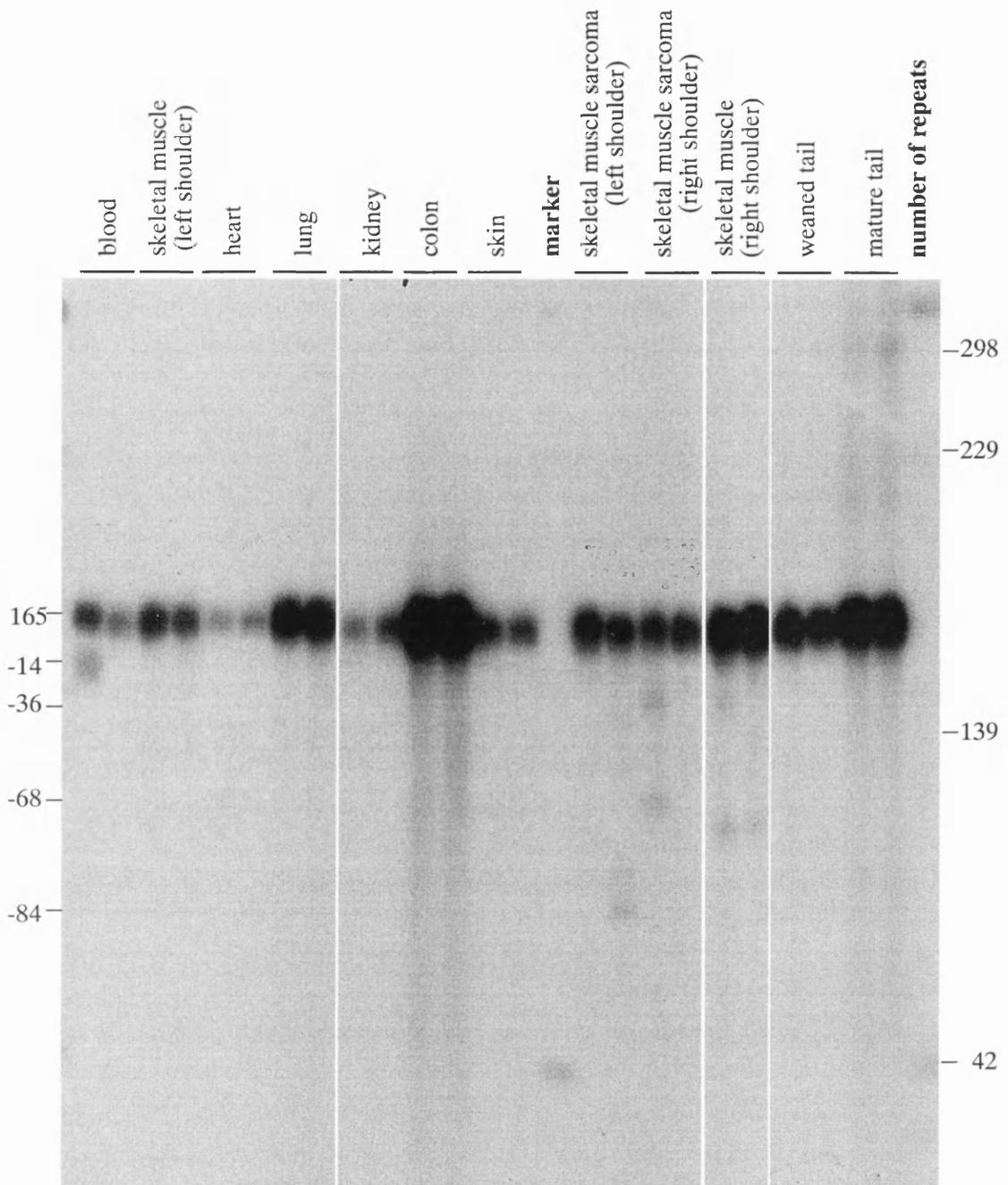
Figure 6.3

Somatic mosaicism in a 12 month old *Dmt*-E mouse with multiple tumours
DNA from a 12 month old *Dmt*-E female with a repeat size of 160 was extracted from multiple tissues including muscle sarcomas and the muscles associated with the tumours. Between 1 and 20 molecules of DNA were used as the template for duplicate SP-PCR amplifications, the PCR products were resolved on a 1.25 % agarose gel (w/v) before transfer to a nylon membrane and Southern hybridisation with DM56. The marker has been converted to show the number of repeats. The size of deletions have also been indicated.
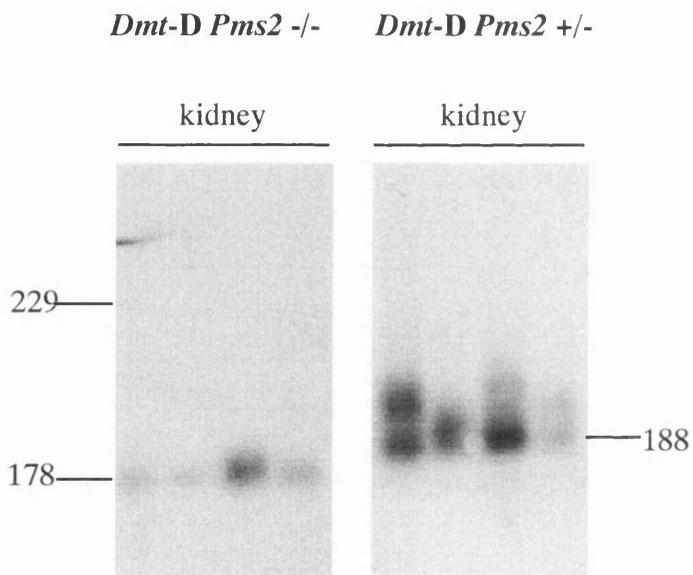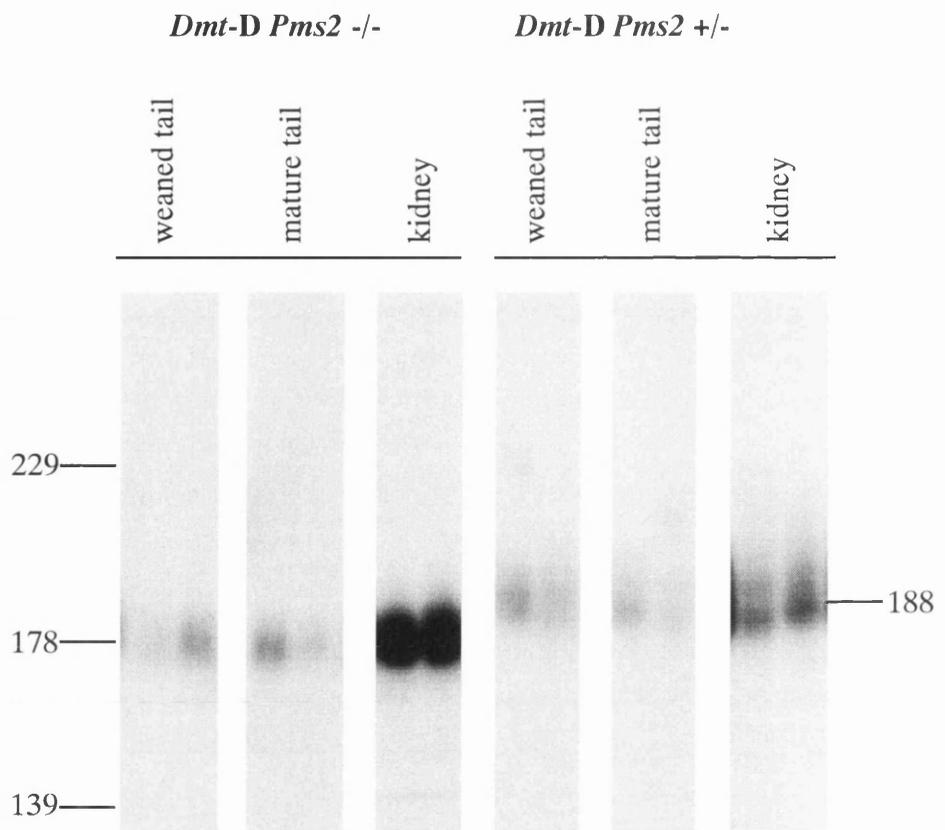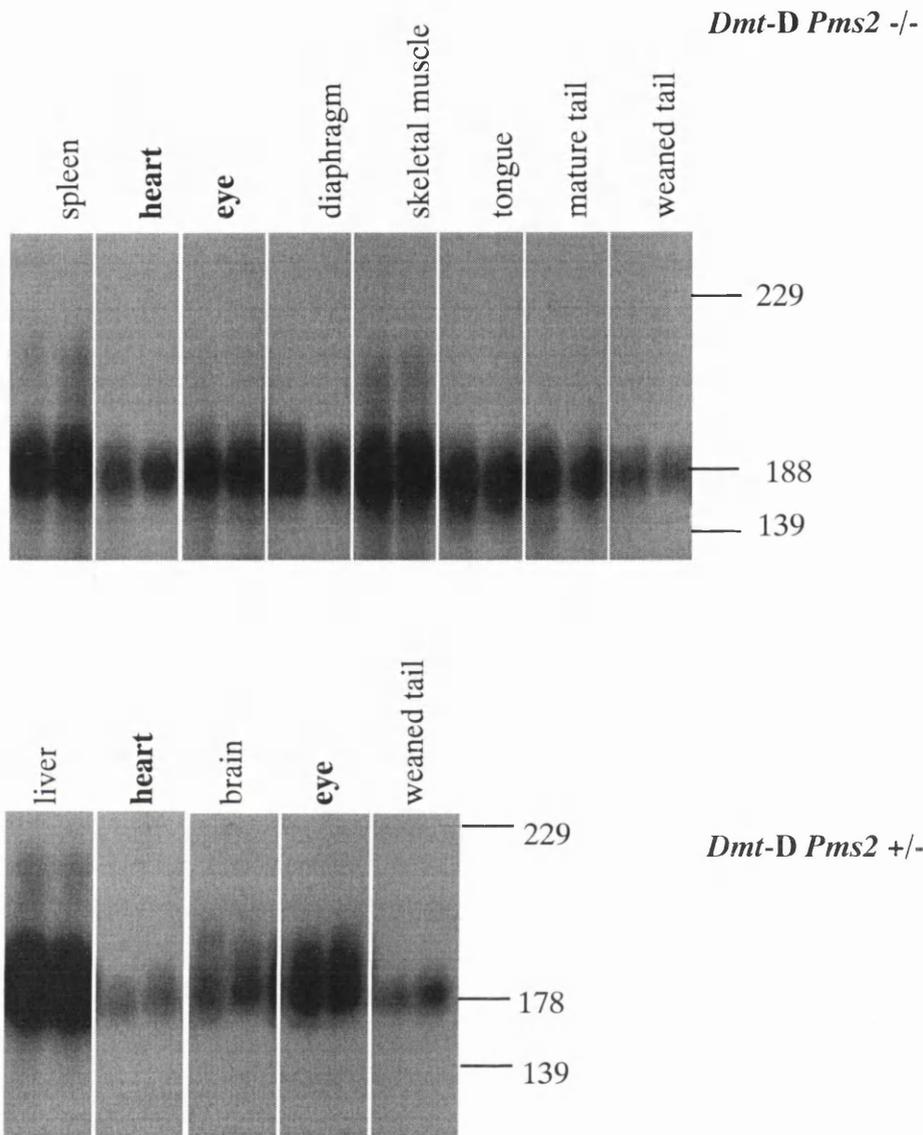
Figure 6.4
Comparision of tissue instability in Dmt-D *Pms2* -/- and Dmt-D *Pms2* +/- littermates
Comparision of tissue instability between a *Dmt*-D *Pms2* +/-mouse with a repeat size of 188 and its *Dmt*-D *Pms2* -/- littermate with a repeat size of 178. Both mice were sacrificed at 6 months. Small pool PCR was carried out on between 1 and 20 molecules of template DNA /reaction. The samples were resolved on a 1.25% agarose gel (w/v) before transfer to nylon membrane and hybridisation with DM56. The marker has been indicated as number of repeats. Highlighted are the common tissues, brain and eye.

6.6 Discussion

The hypothesis of the mismatch repair model for triplet repeat instability suggests that the mismatch repair complex provides the mechanism responsible for the expansion biased instability observed at microsatellite loci and expanded triplet repeat loci (Figure 3.1). This suggests that the loss of the mismatch repair complex could confer stability to regions of instability. This series of experiments was a preliminary investigation of the behaviour of the *Dmt* transgenic repeat on a *Pms2* deficient background. The loss of *Pms2* appears to have little effect on the stable line *Dmt*-E, except in tissues that have undergone a mutation event that has lead to tumour development. Such tissues exhibit a mosaicism in the repeat length and this instability is deletion biased. It should be noted that *Dmt*-E mice not lacking *Pms2* have also exhibited instability after the development of tumours (3.4.4), although repeat deletions and expansions were observed in that case. Pereira and Ingram, (personal communication) also report that quantification of tissues from 8 month old *Dmt*-D *Pms2* -/- and *Dmt*-D *Pms2* +/- littermates, has revealed fewer expansions in the *Dmt*-D *Pms2* -/- mouse although large deletion events appear to occur more frequently in the tissues of these mice. It could be that some mutation events leading to the development of tumours might also effect the mechanisms of repeat instability. Schumacher *et al* (1997) suggests the involvement of MMR in at least 2 pathways involved with repeat stability and it is already known that instability events in triplet repeat tracts contained in plasmids are influenced by the purity and length of the repeat tract (Jaworski *et al.*, 1995) and the loss of MMR proteins. Repeat tracts in excess of 100 almost always result in deletion events greater than 8 repeats (Parniewski *et al.*, 2000), while observations of tracts made up of less than 100 repeats in found instability event of (+/-) one repeat more frequent. The loss of MMR proteins resulted in a reduction of instability events. The purity of the repeat tract also has an effect with purer repeat tracts being more vulnerable to repeat deletions. As all of the roles of repair complex proteins are not known, as *cis* acting factors vary depending on the transgene integration sites and as the purity of the repeat tract has not been discerned in either the lines or individual mice, the processes involved in these repeat events cannot be discerned. That some events involving tumour development do influence repeat instability can be inferred.

Preliminary comparison of the tissues of *Dmt*-D, *Pms2* +/- and *Dmt*-D, *Pms2* -/- siblings suggests that mice completely lacking in *Pms2* exhibit a different somatic instability profile to its *Pms2* heterozygous sibling and were less prone to expansions events. The effects due to the loss of a functioning MMR protein complex because of its lack of *Pms2* requires further

investigation. The use of older sib pairs may confirm the trend of repeat instability. Further comparisons with of *Pms2* null homozygotes and heterozygote siblings with *Pms2* homozygote siblings may reveal differences in the instability profiles of all three. Does a partially functioning MMR complex confer the same level of instability to the *Dmt*-D repeat tract as the *Pms2* homozygote? The quantification of such differences is also required. Studies by Manley *et al*, (1999), studied the effect on the somatic tissues of the Huntington's disease triplet repeat expansion model mice generated by Mangiarini *et al*, (1997), when they were crossed with mice deficient in *Msh2* (Andrew *et al.*, 1999). The MSH2 protein is present in both of the *MutS* heterodimers of the MMR complex (Buermeyer *et al.*, 1999). Tissue instability profiles were generated using Genscan® analysis. Comparison of the instability profiles of the CAG repeat in the striatum and heart of *Msh2* +/+ mice with *Msh2* -/- mice finds the range of repeat instability shown in *Msh2* -/- mice to be less than that exhibited by mice with a functioning MMR complex. This in combination with the preliminary *Pms2* instability data supports the hypothesis of the role of the MMR complex in triplet repeat instability. HNPCC kindreds are most likely to be associated with *MLH1* and *MSH2* mutations with fewer than 10 % associated with *PMS2* mutations leading to suggestions that the roles of *MLH1* and *MSH2* in MMR are critical, while the role of *PMS2*, although important may be carried out through alternative pathways.

Chapter 7

Discussion and future studies

7.1 Introduction

The initial aim of this study was to establish the suitability of the lines of mice generated with the *Dmt* transgene as models of triplet repeat instability. It was established that the triplet repeat in one of the lines, *Dmt*-D, models the gross, tissue specific, somatic instability observed in DM1 patients, being both expansion biased and increasing over the lifetime of the mouse. The other lines, *Dmt*-B, *Dmt*-C and *Dmt*-E exhibited only limited instability. As the lines are all derived from the *Dmt* transgene and in the absence of evidence of triplet tract rearrangement transgenic repeat stability differences must be explained by differences in the integration sites. Analysis of transmission data found that the size of the repeat transmitted was dependent on the sex of the transmitted parent, with males most likely to transmit expansions and females deletions in both of the lines studied, *Dmt*-D and *Dmt*-E. *Dmt*-D and the stable lines were used to investigate candidate mechanisms of expansion in the first instance.

7.2 Mechanisms of somatic instability

Initial investigation of the levels of somatic instability in the different tissues, revealed there to be no correlation between the level of instability exhibited by a tissue and the rate of cell turnover in that tissue. No relationship between the degree of repeat instability shown and tissues that exhibit DM1 disease phenotype was observed. The sex of the mouse had no effect on the somatic instability profile of the tissues examined.

Quantification of the repeat length in the unstable *Dmt*-D tissue, the kidney, revealed that the repeat lengths split into at least three modes as the mouse aged beyond a year. Microdissection of the kidney, an organ containing many specialised cell types, suggests that the cell population in different regions are subject to different levels of repeat expansion events. So different cell populations, both between and within the same tissue, subject the repeat tract to

145

differing numbers of repeat instability events. Comparison of the rate of replication and the level of transcription of the different tissues in *Dmt*-D found no obvious correlation. It may be the case that replication and / or transcription may result in repeat instability events, but not at a level that accounts for the inter and intra cell differences observed or, maybe replication and / or transcription are necessary to the triplet repeat instability pathway, but not sufficient alone to promote instability at the observed levels.

Although *PMS2* is not reported as the most important component of the *MutLα* complex, preliminary studies at the 6 month timepoint appears to reduce the level of expandability exhibited by somatic tissue. Subsequent studies carried out by Mr Mario Pereira and Miss Laura Ingram, an honours student, have confirmed that the repeat is more stable in a *Dmt*-D *Pms2* -/- mouse than in its *Dmt*-D *Pms2* +/- littermate when studying tissues of eight month old mice. However, initial analysis indicates that the differences in stability between the tissues of the littermates shows a tissue specific pattern.

The MSH2 protein is present in both of the *MutS* complexes of MMR (Buermeyer *et al.*, 1999). Mice containing the triplet repeat expansion in the Huntington's disease transgenic mice generated by Mangiarini *et al*, (1997), were crossed with mice deficient in *Msh2* (Andrew *et al.*, 1999) by Manley *et al*, (1999). Genscan® analysis was used to study the effect of the loss of *Msh2* on the transgenic repeat instability exhibited by different tissues. Comparison of the instability profiles of the CAG repeat in the striatum and heart of *Msh2* +/+ mice with *Msh2* -/- mice finds the range of repeat expandability exhibited by the *Msh2* -/- mice to be less than that exhibited by mice with a functioning MMR complex. These data and the preliminary *Pms2* studies of the *Dmt*-D mice suggest that the MMR protein complex has a role in the expansion mechanism(s).

There are cells contained within the kidney which undergo a higher number of expansion biased repeat instability events than any of the other cells which have been studied. The mechanism of expansion involves a functioning MMR complex, but transcription and or replication events are not sufficient to account for the majority of instability events, although a role, possibly essential, has not been excluded.

The kidney contains highly specialised cells and as none of the other tissues, with which it has cell types in common, exhibit the extreme level of repeat expansion it is likely that the unstable cell type is among these specialised cells. My preliminary microdissection of kidney regions would suggest this is so. It has been observed that oxidative stress, a condition describing the production of oxygen radicals beyond a threshold for proper antioxidant

146

neutralisation has been implicated in the pathological mechanism of several neurodegenerative disorders, including Huntington's disease (Borlongan *et al.*, 1996). A study by Usuki and Ishiura (1998) on the effect of oxidative stress on a cell line containing *DMPK* with an expanded CTG repeat (CTG)46 observed that the *DMPK* expanded repeat increased the cells susceptibility to oxidative stress. Usuki and Ishiura (1998) suggest that the expansion may in some disrupt the pathway controlling signalling molecules effecting the redox state.

I hypothesise that the cell type which is exhibiting the high level of repeat expansion events, is subject to a high level of oxidative stress. This damages the DNA that requires the action of the MMR mechanism to repair it. A functioning mismatch repair protein complex can misinterpret the repeat strands and instability events occur. An excess of oxygen radicals in these cells damage the DNA, including the areas of the *Dmt* repeat expansion, which reanneal out of register after the DNA has been repaired, in a similar manner to that outlined in Figure 5.1. The functioning MMR system misinterprets the misaligned repeats and repairs them, resulting in a change in the repeat size. A threshold for reach instability events of around 100 repeats has been hypothesised, beyond this threshold, repeat instability events are more likely to be expansions (Schumacher *et al.*, 1997). The more oxidative stress a cell is subjected to, the more likely it is that the DNA is damaged and requires repair.

That the *Dmt*-D repeat is more vulnerable to these events than the other lines then I further hypothesise that the context of the *Dmt*-D integration site is more vulnerable to damage by oxidative stress than the integration sites of the other lines. For example if *Dmt*-D has integrated into a gene rich region it will be regularly accessible for transcription, but also vulnerable to damage which therefore requires repair.

7.3 Transmission studies

The size of the repeat expansion and the sex of the parent it is transmitted by, influences the size of the repeat transmitted to a DM1 patient (Lavedan *et al.*, 1993b). Parental age has been confirmed to influence the size of repeat in Huntington's and DRPLA transmissions (Duyao *et al.*, 1993; Sato *et al.*, 1999). Observation of transmission patterns of *Dmt*-D and *Dmt*-E has seen a sex dependent difference in the pattern, with the males of both lines being more likely to transmit an expansion than the females which are most likely to transmit deletions. Despite each line carrying almost identical transgenes, the variation in transmitted repeat length is greater in the offspring of *Dmt*-D mice underline that factors other than parental sex are influencing the size of the repeat observed in the offspring.

No correlation was observed between the age of the transmitting parent and the size of the repeat transmitted. Such a correlation was observed in male transmissions of DM-300 DM1 model mice outlining a role for the genomic DNA flanking the DM1 repeat as these transgenic models contain an expanded DM1 repeat within 45kb of outlying DNA, including *DMPK, SIX5* and *DMWD*.

The observation that the sex of the offspring of the *Dmt*-E males influences the size of the repeat that it carries, suggests the influence of the sex chromosome within the gamete on the repeat. Perhaps the influences a required repair event prior to fertilisation or a role for somatic instability in the initial stages of the zygotes development, although the mechanism involved, and the effect that any initial germline instability may have on subsequent events is unknown. That the pattern of segregation distortion observed in the *Dmt*-D and *Dmt*-E male and female transmissions was dependent on the sex and the line indicates firstly that there are sex differences effecting the size of the repeat transmitted to the offspring and secondly the effect of the integration site of the transgene on the repeat size transmitted to the offspring.

7.4 DM1 transgenic mice models

When this work was begun it had been established that the *Dmt* mice generated by Monckton *et al.* (1997), from DM1 patient genomic DNA consisting of a (CTG)162 tract flanked by ~800 bp of DNA, exhibited intergenerational instability and limited somatic instability (Figure 1.3). Another DM1 derived transgene, (DM 55) was generated by Gourdon *et al.* (1997), (Figure 1.3). This transgene consisted of a (CTG)55 tract contained within ~ 45 kb of sequence including the coding sequence for *DMPK, DMWD* and, *SIX5*. As with the *Dmt* transgenic lines, the (DM 55) lines also integrated randomly into the genome. Unlike the *Dmt* transgenic lines the effect of the murine integration site was less likely to influence the transgenic repeat because of the amount of flanking DNA. Of the seven DM55 transgenic lines generated, six of the lines exhibited gametic instability and they contained between one and four copies of the transgene at a single integration site. Of the total progeny studied, 6.8% exhibited a repeat size differing to that of its transmitting parent. A more in depth analysis of the DM55 DM1 murine models was carried out by Lia *et al.* (1998). Inter tissue comparison of the repeat in the lung, kidney, pancreas, liver, ovary, eye, skeletal muscle (quadriceps), gastronemius and heart in individuals aged between 4 and 19 months from the two single copy transgenic lines, DM55/86 and DM55/85 was carried out. Analysis revealed tissue specific, expansion biased, instability that increased with the age of the mouse. The order of instability

that the studied tissues could be placed in, i.e. liver and kidney most unstable, lung most stable, was similar between the lines. However, the instability profile of the tissues was not identical, despite the large amount of transgenic DNA, which flanks the repeat and the fact that both lines have been generated on the same inbred strain. That the level of instability presented by each tissue is not identical between the two lines is a reflection that DNA outlying the transgene is also having an effect. It is possible that *DMWD* and *SIX5* are not the only genes in *cis* to *DMPK* that may influence the repeat instability. Other genes have been reported in the polygenic region around *DMPK*, (Alwazzan *et al.*, 1998). Within the 300kb that include *DMPK*, *SYMPLEKIN*, a polypeptide encoded by 1147 aa is found upstream of *SIX5* while two further genes were found downstream of *DMPK*, the *GASTRIC INHIBITOR POLYPEPTIDE RECEPTOR GENE (GIPR)* and *20D7*, which is expressed predominantly in the testis. These genes may have roles in repeat tract instability and or DM1 phenotypes. The DM55 transgene did not include these genes. The differences in the integration sites of DM55/85 and DM55/86 are still affecting the behaviour of the repeat despite the large amount of flanking DNA.

Quantification of the level of instability observed in DM55/86 in a 4 month and a 19 month old individual using small pool PCR analysis confirmed the somatic instability to be expansion biased. In common with the observations of *Dmt*-D the heart was observed to be particularly stable. The muscle and eye also had instability profiles similar to those observed in *Dmt*-D. The kidney and the liver of 19 month old DM55/86 were the most unstable tissues studied with expansions of ~+15 from the progenitor repeat size. The kidney of *Dmt*-D also exhibits the greatest instability profile although the expansions observed are greater. These data probably indicate a balance between the effect of the flanking DNA and the effect of the number of repeats.

To assess whether the level of transgenic transcription might be influencing the somatic instability observed in the DM55 lines, semi quantitative RT-PCR analysis was carried out to assess the level of transcription of transgenic *DMWD*, *DMPK* and *SIX5* in the brain, heart, muscle and kidney. Transcripts of all three genes were found in all seven lines. Moreover the tissue specific expression pattern observed in *DMWD* and *DMPK* was conserved in the DM55 mice. As all of the DM55 lines exhibited gametic or somatic instability, (no gametic instability had been observed in the somatically unstable DM55/85), Lia *et al*, 1998 feel transcription of the repeat may be instrumental to the instability observed, although the most unstable tissues, the kidney and the liver, did not undergo the highest level of transcription.

As with the *Dmt*-D line, there was no correlation between the level of cell turn over ((Messier and Leblond, 1960), summarised in table 3.2) and the degree of repeat instability observed in each tissue. The tissues associated with the DM1 disease phenotype, i.e. muscle, eye and heart did not exhibit instability greater than unaffected tissues such as the kidney and liver. This observation is also in common with *Dmt*-D. The lack of availability of DM1 patient tissue means that a parallel observation has not been made in DM1 patients. However, the somatic studies of the *Dmt* and DM55 lines may suggest that a similar observation will be observed in DM1 patients.

To further investigate the effect of the flanking DNA of DM1 patients, and how it may relate to the length of the repeat tract, transgenic lines additional to DM55 were generated (Seznec *et al.*, 2000). The DNA used in the transgenes DM20 and DM300 originated from people in the same family as the patient who contributed the DM55 DNA. Studies found neither gametic or somatic instability in the DM20 lines. In contrast transmission studies of the DM300 lines revealed that a repeat larger than that of the transmitting parents was passed on in over 86% of the transmissions. Comparison of a range of tissues found them to exhibit somatic instability that varied from tissue to tissue. Again the order of instability of the tissues appeared to be in common with that observed in DM55/86 and DM55/85. The most unstable tissue, liver exhibited an expansion of ~100 repeats.

The difference in both the gametic and somatic stability levels observed in the DM20, DM55 and DM300 lines parallels the thresholds of repeat stability observed in DM1 patients, suggesting when the repeat is contained within a similar genomic context, the repeat expansion thresholds are similar in both humans and mice (Seznec *et al.*, 2000).

Analysis of the DM300 transmissions found that although the transmissions were expansion biased, the larger expansions were transmitted paternally. Indeed when the size of the repeat transmitted was compared to the age of the parent at conception, the size of paternal transmissions increased as the age of the mouse increased. No similar effect was observed in maternal transmissions.

7.5 Complementary studies of the *Dmt* transgenic mice

Examination of the effect of *cis* acting factors is limited in *Dmt* lines. Dr Graham Brock of the University of Glasgow used vectorette PCR to isolate genomic sequence flanking *Dmt*-D and *Dmt*-E and to determine if possible the integration sites. The location of *Dmt*-E is unlikely to be elucidated due to its integration into a LINE-1 repetitive element, a motif found throughout the murine genome. *Dmt*-D has integrated into a novel region of chromosome 11. Although this region has not yet been sequenced as part of the murine sequencing project, the location *Dmt*-D will be identified as more murine genome is sequenced. While this in itself may not explain the instability pattern observed in the tissues of this line, information about the sort of DNA that *Dmt*-D has integrated into, and the nature of the gene which is causing *Dmt*-D to be transcribed may contribute to the understanding of the expansion mechanism.

The triplet repeat of DM1 is located within a CpG island. Dr Graham Brock investigated the methylation pattern of the 609 bp region of the transgene using bisulfite induced modification of genomic DNA, a reaction which will convert cytosine to uracil but does not effect 5-methyl cytosine (Brock and Monckton, 2000). The *Dmt*-E and *Dmt*-D transgenic lines in a number of tissues and timepoints were studied. A line specific difference was observed. In the first instance 18 month old mice were studied. The region flanking *Dmt*-D was found to become methylated over time in a tissue specific manner. The same tissues in *Dmt*-E mice had not become methylated at this location. Comparison of *Dmt*-D and *Dmt*-E mice at 2 months again found no methylation in the *Dmt*-E mice but partial methylation in the tissues of young *Dmt*-D mice, suggesting that the methylation pattern develops over time. An in depth examination of different tissues in *Dmt*-D mice revealed different level of methylation in different tissues at the same timepoint. Only the brain of *Dmt*-E revealed any change in the methylation pattern and this was partial. That *Dmt*-D does not appear to be methylated in young mice suggests that the methylation pattern of the *Dmt*-D transgene does not effect the stability pattern of *Dmt*-D. The methylation patterns observed do exhibit differences between the tissues. We hypothesise however, that the instability of the repeat tract in some way induces the methylation of the flanking region.

A number of *Dmt*-D tissues from a selection of timepoints have been used to establish cell lines by Mr Mario Pereira (Gomes-Pereira *et al.*, 2001). The lines studied include ones derived from the eye, lung and kidney of a 6 month old male. Study of these cell lines over time has found the accumulation of multiple small expansion biased mutations observed *in vivo* in *Dmt*-D was also observed in the *in vitro* model. The repeat being particularly unstable in kidney

cells. Observation of the cell lines over hundreds of days and passages found the repeat in the lung cell line to be particularly stable. In contrast the cultured kidney cells were particularly unstable. Suprisingly selection for larger alleles appeared to take place within the first few passages, that is some sort of bottleneck occurred and a mean increase from the progenitor repeat size recorded. These observations again suggest that different cell populations within a tissue type were subject to different levels of repeat expansion events.

7.6 Suggestions for future experiments

Although the *Dmt*-D model has been shown to be a suitable model for the unstable repeat in DM1, it has not lead to the elucidation of the mechanism of repeat expansion. The mechanism that leads to the somatic instability of an expanded repeat tract will probably be multifactorial. This makes the investigation of expansion mechanisms difficult. However, the *Dmt*-D transgenic line provides a DM1 model line on an isogenic background. The only differences between individuals being their repeat length. The controlled introduction of variants or the investigation of spontaneously occurring modifications / mutations may in some way explain the complexity of the instability process:

- The investigation of tumours occurring in the *Dmt* lines and their effect on the transgenic repeat if the nature of the mutations that have lead to the development of the tumour may be identified.

- Both *Dmt*-D and *Dmt*-E mice are currently being bred onto different inbred mouse backgrounds. The BalB C, Black 6 and C3 H mouse strains have been chosen because of their diversity from the FVB/n strain and each other. The level of repeat instability and how this relates to the disease phenotype of triplet repeat diseases can vary even between family members carrying similar repeat sizes. The diverse backgrounds of DM1 patients may suggest that polymorphisms that may result in simple and normally inconsequential variation in protein function or efficiency, may in individuals carrying a pre mutation or a disease range repeat expansion at a triplet repeat disease locus result in instability at the repeat loci, resulting in RNA transcripts which in someway disrupt cell function and leads to the disease phenotype. The introduction of the *Dmt* repeat onto the different strain backgrounds will expose the repeat to different *trans* acting factors. Examination of the tissue instability on these different backgrounds may reveal differences, which may be traced to polymorphisms in particular proteins. For example mutations to proteins that form part of the MMR complex, which may have subtle effects on the efficiency of repair.

- It is also possible that the mice may present reduced aggression levels on a different strain background and that the incidence of successful matings may increase as may the efficiency of subsequent superovulation.

- Murine knockouts for *Mlh1* and *Msh2* genes that code for other proteins in the MMR complex also exist. The effect of the loss of these gene products on *Dmt*-D repeat instability may also allow further elucidation of the effect of MMR on repeat instability (Andrew *et al.*, 1997; Baker *et al.*, 1996). Further characterisation of the effect of a deficient *Pms2* background on the *Dmt* repeat instability levels should be carried out.

- I have hypothesised that the cell population in kidney that results in the dramatic expansions are subject to oxidative stress, that this results in a level of DNA damage unrivalled in other cell populations. To identify regions of the kidney that are subject to oxidative stress transverse sections of both wildtype and *Dmt*-D kidneys should be isolated and the level of oxidative stress that different regions are subject to detection using an antibody to a marker of oxidative stress, i.e. superoxidase dismutase. If this correlates with regions in which the highest level of repeat instability occur then they may be a connection between cells subjected to high levels of oxidative stress and high levels of repeat instability.

- The effect of other mechanisms of repair on the *Dmt*-D expanded repeat tract could be investigated. Mutants of genes involved in transcription coupled nucleotide excision repair may also be found to influence repeat stability levels.

- The isolation of mature sperm cells of *Dmt*-E mice and the elucidation of the repeat size and the sex chromosome carried by that cell may narrow the window in which repeat instability events occur which are influenced by the offsprings sex.

7.7 Conclusions

The mechanism involved in the expansion of triplet repeats is **multifactorial,** as is the level of instability a repeat tract will exhibit.

The mechanism of actual triplet repeat expansion is in some way related to the MMR mechanism. While the nature of the instability event(s) that the repeat tract experiences, whether there is an expansion or a deletion, may depend on the length of the repeat. There is a threshold of repeat length beyond which a repeat tract becomes more prone to expansions (Schumacher *et al.*, 1997).

153

- There was no obvious correlation between the level of tissue replication and the level of instability found in that tissue.
- There was no obvious correlation between the level a tissue was transcribed and the level of instability found in that tissue.
- The role that transcription and or replication may play in triplet repeat instability has not been determined, just that neither relates directly to the gross level of triplet repeat tract instability observed in *Dmt*-D. Transcription and or replication may play a role essential in the instability event pathway, but this has not been determined in this series of experiments.
- That *Dmt*-D exhibits tissue specific repeat instability that is not observed in the other lines indicates the importance of the genomic context in which the repeat tract lies. For example if the chromatin structure is such that the repeat may be transcribed at a high level, the DNA may be vulnerable to other factors such as free radical damage. The repair of this damage may result in repeat instability events. Continuing levels of transcription might result in further DNA damage and repeat instability events.
- If a particular cell is subject to events that may damage its DNA, it will be subject to DNA repair mechanisms more often. Thus a repeat tract, if present will be subject to multiple small changes in repeat length, more often than other cells, which may be located in the same tissue, but due to their role undergo less damage to their DNA and are thus subject to fewer DNA repair events.
- Gametic instability follows different patterns in each of the sexes; Male and female gametes follow different maturation pathways with the oocyte perhaps being more vulnerable to DNA damaging events as it is present for the lifetime of the mouse. However, the DNA that flanks the integration site of the *Dmt*-D and *Dmt*-E transgenes influence the size of the transmitted repeat.

This series of experiments has examined the somatic instability in a range of tissues in stable and unstable triplet repeat expansion models. That *Dmt*-D does model tissue specific expansion biased instability suggests that this is a model suitable for addressing questions posed by DM1 and other expanded triplet repeat diseases. It is known that DM1 exhibits different levels of somatic mosaicism in different tissues, as revealed by the leukocyte / muscle studies of Monckton *et al.* (1995) and also a regional variation within the brain (Ishii *et al.*, 1996), which is in common with the neurodegenerative triplet repeat diseases. What is not

known, due to the limited availability of archived repeat disease tissue is how the repeat instability profile varies from tissue to tissue and over time. These experiments have demonstrated that even the most stable tissues exhibit a degree of expansion biased instability over the lifetime of the mouse and the most unstable tissues studied do not correlate with the tissues that exhibit a disease phenotype in DM1, although it should be noted that *Dmt* mice do not model DM1 symptoms. However, this observation may suggest that the expansion alone is not sufficient to cause the disease phenotype but that it is a combination of the role of the cell in which the expansion occurs and the expansion which result in the disease phenotype. This sort of idea moves forward the thinking around triplet repeat diseases, which without murine models would still be confined to archive tissues.

A great deal has mean made in this thesis about elucidating the mechanism of repeat expansion, but the reasons for doing this have not been outlined. The correlation between the triplet repeat expansions and disease phenotypes is well known, as is the incidence of these diseases within family pedigrees. If the mechanisms of expansion are understood then this allows the development of therapies targeting these molecular mechanisms. Current treatments for these diseases involve the treatment of the symptoms.

Finally, the *Dmt* mice, most specifically the *Dmt*-D mice, provide a model for studying other genetic effects which may be too subtle to detect within diverse human populations, but upon their characterisation may be found to occur within humans. The *Dmt*-D transgene is currently presented on an isogenic background. The study of this transgene on other isogenic backgrounds may result in variations in the tissue instability profiles presented. These variations would be accounted for by the genetic differences between the strains, identification of these factors may lead to the identification of subtle but key factors that effect not only repeat instability but other mutations events, for example cancer.

References

Akiyama, Y., Sato, H., Yamada, T., Nagasaki, H. and A, T. (1997). Germ-line mutation of the hMSH6/GTBP gene in an atypical hereitary nonpolyposis colorectal cancer kindred. *Cancer Research* **57**, 3920-3923.

Alwazzan, M., Hamshere, M. G., Lennon, G. G. and Brook, J. D. (1998). Six transcripts map within 200 kilobases of the myotonic dystrophy expanded repeat. *Mammalian Genome* **9**, 485-7.

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. *Human Molecular Genetics* **8**, 1491-1497.

Andrew, S. E., McKinnon, M., Cheng, B. S., Francis, A. and Penney, J. (1999). Tissues of MSH2 deficient mice demonstrate hypermutability on exposure to a DNA methylating agent. *Proceedings Of the National Academy Of Sciences Of the United States Of America* **95**, 1126-30.

Andrew, S. E., Reitmair, A. H., Fox, J., Hsiao, L., Francis, A., McKinnon, M., Mak, T. W. and Jirik, F. R. (1997). Base transitions dominate the mutational spectrum of a transgenic reporter gene in MSH2 deficient mice. *Oncogene* **15**, 123-9.

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. *Human Molecular Genetics* **2**, 1397-1400.

Ashizawa, T., Dubel, J. R., Tran, L., Harati, Y., Perryman, M. B. and Epstein, H. F. (1993). Muscle DNA shows different triplet expansion size from peripheral blood leukocyte DNA in myotonic dystrophy. *Neurology* **43**, 2674-2678.

Ashizawa, T., Dunne, P. W., Ward, P. A., Seltzer, W. K. and Richards, C. S. (1994). Effects of sex of myotonic dystrophy patients on the unstable triplet repeat in their affected offspring. *Neurology* **44**, 120-122.

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

Baker, S., Bronner, C. E., Zhang, L., Plug, A. W., Robatzek, M., Warren, G., Elliott, E. A., Yu, J., Ashley, T., Arnheim, N., Flavell, R. A. and Liskay, R. M. (1995). Male mice defective in the DNA mismatch repair gene *PMS2* exhibit abnormal chromosome synapsis in meiosis. *Cell* **82,** 309-319.

Baker, S. M., Plug, A. W., Prolla, T. A., Bronner, C. E., Harris, A. C., Yao, X., Christie, D. M., Monell, C., Arnheim, N., Bradley, A., Ashley, T. and Liskay, R. M. (1996). Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over. *Nature Genetics* **13,** 336-342.

Bates, G., Mangiarini, L., Mahal, A. and Davies, S. (1997). Transgenic models of Huntington's disease. *Human Molecular Genetics* **6,** 1633-1637.

Batten, F. E. and Gibb, H. P. (1909). Myotonia atrophica. *Brain* **32,** 187-205.

Bell, J. (1948). Dystrophia myotonica and allied diseases. In *Treasury of Human Inheritance.* Cambridge: Cambridge University Press.

Benders, A. A. G. M., Groenen, P. J. T. A., Oerlemans, F. T. J. J., Veerkamp, J. H. and Wieringa, B. (1997). Myotonic dystrophy protein kinase is involved in the modulation of the Ca2+ homeostasis in skeletal muscle cells. *Journal Of Clinical Investigation* **100,** 1440-1447.

Berul, C. I., Maguire, C. T., Aronovitz, M. J., Greenwood, J., Miller, C., Gehrmann, J., Housman, D., Medelsohn, M. E. and Reddy, S. (1999). DMPK dosage alterations results in atrioventricular conduction abnormalities in a mouse myotonoc dystrophy model. *The Journal of Clinical Investigation* **103,** R1-R7.

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

Borlongan, C., Kanning, K., Poulos, S., Freeman, T., Cahill, D. and Sanberg, P. (1996). Free Radical Damage and Oxiditve Stress in Hungtington's Disease. *Journal of the Florida Medical Association* **83,** 335-341.

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

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. *Human Molecular Genetics* **8,** 1061-1067.

Brock, G. J. and Monckton, D. G. (2000). *Cis*-acting modifiers of somatic repeat instability in a transgenic mouse model of a triplet repeat expansion: the role of methylation. In *Tenovus 2000*. Glasgow.

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., Sohn, R., Zemelman, B., Snell, R. G., Rundle, S. A., Crow, S., Davies, J., Shelbourne, P., Buxton, J., Jones, C., Juvonen, V., Johnson, K., Harper, P. S., Shaw, D. J. and Housman, D. E. (1992). Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **68,** 799-808.

Brunner, H. G., Bruggenwirth, H. T., Nillesen, W., Jansen, G., Hamel, B. C. J., Hoppe, R. L. E., de Die, C. E. M., Howeler, C. J., van Oost, B. A., Wieringa, B., Ropers, H. H. and Smeets, H. J. M. (1993). Influence of sex of the transmitting parent as well as of parental allele size on the CTG expansion in myotonic dystrophy (DM). *American Journal Of Human Genetics* **53,** 1016-1023.

Brunner, H. G., Nillesen, W., van Oost, B. A., Jansen, G., Wieringa, B., Ropers, H. H. and Smeets, H. J. M. (1992). Presymptomatic diagnosis of myotonic dystrophy. *Journal Of Medical Genetics* **29,** 780-784.

Buermeyer, A. B., Wilson-Van Patten, C., Baker, S. M. and Liskay, R. M. (1999). The human MLH1 cDNA complements DNA mismatch repair defects in Mlh1- deficient mouse embryonic fibroblasts. *Cancer Research* **59,** 538-41.

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

Buxton, J., Shelbourne, P., Davies, J., Jones, C., Van Tongeren, T., Aslanidis, C., de Jong, P., Jansen, G., Anvret, M., Riley, B., Williamson, R. and Johnson, K. (1992). Detection Of an Unstable Fragment Of DNA Specific to Individuals With Myotonic Dystrophy. *Nature* **355,** 547-548.

Cameron, I. L. (1970). Cell renewal in the organ and tissues of the nongrowing adult mouse. *Tex Reproductive Biological Medicine* **28,** 203-248.

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

Carey, N., Johnson, K., Nokelainen, P., Peltonen, L., Savontaus, M. L., Juvonen, V., Anvret, M., Grandell, U., Chotai, K., Robertson, E., Middletonprice, H. and Malcolm, S. (1994). Meiotic Drive At the Myotonic Dystrophy Locus. *Nature Genetics* **6,** 117-118.

Chen, W. and Jinks-Robertson, S. (1999). The role of the mismatch repair machinery in regulating mitotic and meitotic recombination between diverged sequences in yeast. *Genetics* **151,** 1299-1313.

Clark, A. J. H., G, Yull, F E. (1997). Mammalian cDNA and prokaryotic reporter sequences silence adjacent transgenes in transgenic mice. *Nucleic Acids Research* **25,** 1009-1014.

Cummings, C. J. and Zoghbi, H. Y. (2000). Fourteen and counting: unraveling trinucleotide repeat diseases. *Human Molecular Genetics* **9,** 909-16.

David, G. A., N. Stevanin, G. Durr, A, Yvery, G. Cancel, G. Weber, C. Imbert, G. Saudou, F. Antoniou, E. Drabkin, H. Gemmill, R. Gianti, P. Benomar, A. Wood, N. Ruberg, M. Agid, Y. Mandel, JL. Brice, A. (1997). Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nature Genetics* **17,** 65-70.

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

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. *Proceedings Of the National Academy Of Sciences Of the United States Of America* **94,** 7388-7393.

Deka, R., Majumder, P. P., Shriver, M. D., Stivers, D. N., Zhong, Y., Yu, L. M., Barrantes, R., Yin, S. J., Miki, T., Hundrieser, J., Bunker, C. H., McGarvey, S. T., Sakallah, S., Ferrell, R. E. and Chakraborty, R. (1996). Distribution and evolution of CTG repeats at the myotonin protein kinase gene in human populations. *Genome Res* **6,** 142-54.

Duyao, M., Ambrose, C., Myers, R., Novelletto, A., Perischetti, F., Frontali, M., Folstein, S., Ross, C., Franz, M., Abbott, M., Gray, J., Conneally, P., Young, A., Penney, J., Hollingsworth, Z., Shoulson, I., Lazzarini, A., Falek, A., Koroshetz, W., Sax, D., Bird, E., Vonsattel, J., Bonilla, E., Alvir, J., Bickham-Conde, J., Cha, J.-H., Dure, L., Gomez, F., Ramos, M., Sanchez-Ramos, J., Snodgrass, S., de Young, M., Wexler, N., Moscowitz, C., Penchaszadeh, G., MacFarlane, H., Anderson, M., Jenkins, B., Barnes, G., Srinidhi, J., Gusella, J. and MacDonald, M. (1993). Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nature Genetics* **4,** 387-392.

Engelkamp, D. and van Heyningen, V. (1996). Transcription factors in disease. *Current Opinions in Genetic Development* **6,** 334-342.

Erikkson, M., Ansved, T., Edstrom, L., Wells, D. J., Watt, D. J., Anvrett, M. and Carey, N. (2000). Independent regulation of the myotonic dystrophy 1 locus genes postnatally and during adult muscle regenaration. *The Journal of Biological Chemistry* **275,** 19964-19969.

Festing, M. (1998). Inbred strains of mice.
http://www.informatics.jax.org/external/festing/search_form.cgi .

Flynn, G. A., Hirst, M. C., Knight, S. J. L., Macpherson, J. N., Barber, J. C. K., Flannery, A. V., Davies, K. E. and Buckle, V. J. (1993). Identification Of the Fraxe Fragile Site In 2 Families Ascertained For X-Linked Mental-Retardation. *Journal Of Medical Genetics* **30,** 97-100.

Friedrich, T., Kroger, B., Bialojan, S., Lemaire, H. G., Hoffken, H. W., Reuschenbach, P., Otte, M. and Dodt, J. (1993). A Kazal-type inhibitor with thrombin specificity from Rhodnius prolixus. *Journal of Biolological Chemistry* **268,** 16216-16222.

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. and Caskey, C. T. (1993). Decreased expression of myotonin protein-kinase messenger RNA and protein in adult form of myotonic dystrophy. *Science* **260,** 235-238.

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

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

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,** 16793-7.

Gibbs, M., Collick, A., Kelly, R. G. and Jeffreys, A. J. (1993). A Tetranucleotide Repeat Mouse Minisatellite Displaying Substantial Somatic Instability During Early Preimplantation Development. *Genomics* **17,** 121-128.

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., Jirik, F. and Hayden, M. R. (1996). Absence of disease phenotype and intergenerational stability of the CAG repeat in transgenic mice expressing the human Huntington disease transcript. *Human Molecular Genetics* **5,** 177-185.

Gomes-Pereira, M. R., Fortune, M. T. and Monckton, D. G. (2001). Mouse tissue culture models of unstable triplet repeats: in vitro selection for larger alleles, mutational expansion bias and tissue specificity, but no association with cell division rates. *Human Molecular Gentics* **10,** 845-854.

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. *Nature Genetics* **15,** 190-192.

Haberhausen, G., Damian, M. S., Leweke, F. and Muller, U. (1995). Spinocerebellar ataxia type 3 (SCA 3) is geneticaly identical to Machado-Joseph disease (MJD). *Journal of Neurological Science* **132,** 71-75.

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. *Proceedings Of the National Academy Of Sciences Of the United States Of America* **94,** 7394-7399.

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

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

Harper, P. S. (1989). Myotonic Dystrophy: W B Saunders Company.

Harper, P. S. (1998). Myotonic Dystrophy as a Trinucleotide repeat disorder- A clinical perspective. In *Genetic Instabilities and Hereditary Neurological Diseases* (ed. R. D. Wells and S. T. Warren), pp. 115-141: Academic Press.

Harper, P. S., Harley, H. G., Reardon, W. and Shaw, D. J. (1992). Anticipation In Myotonic Dystrophy - New Light On an Old Problem. *American Journal Of Human Genetics* **51,** 10-16.

Heath, S. K., Carne, S., Hoyle, C., Johnson, K.J. and Wells, D.J. (1997). Characterisation of expression of *mDMAHP*, a homeodomain-encoding gene at the murine DM locus. *Human Molecular Genetics* **6,** 651-657.

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., Sharp, A. H., Ananth, U., Seltzer, W. K., Boss, M. A., Vieria-Saecker, A. M., Epplen, J. T., Riess, O., Ross, C. A. and Margolis, R. L. (1999). Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. *Nature Genetics* **23,** 391-392.

Hudson, A. J., Huff, M. W., Wright, C. G., Silver, M. M., Lo, T. C. Y. and Banerjee, D. (1987). The Role Of Insulin Resistance In the Pathogenesis Of Myotonic Muscular Dystrophy. *Brain* **110,** 469-488.

Huntington's Disease Collaborarive Research Group. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* **72,** 971-983.

Hurst, G. D. D., Hurst, L. D. and Barrett, J. A. (1995). Meiotic Drive and Myotonic Dystrophy. *Nature Genetics* **10,** 132-133.

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*. *Nature Genetics* **13,** 196-202.

Imbert, G., Kretz, C., Johnson, K. and Mandel, J.-L. (1993). Origin of the expansion mutation In myotonic dystrophy. *Nature Genetics* **4,** 72-76.

Imbert, G., Saudou, F., Yvert, G., Devys, D., Trottier, Y., Garnier, J. M., Weber, C., Mandel, J. L., Cancel, G., Abbas, N., Durr, A., Didierjean, O., Stevanin, G., Agid, Y. and Brice, A. (1996). Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nature Genetics* **14,** 285-91.

Ishii, S., Nishio, T., Sunohara, N., Yoshihara, T., Takemura, K., Hikiji, K., Tsujino, S. and Sakuragawa, N. (1996). Small increase in triplet repeat length of cerebellum from patients with myotonic dystrophy. *Human Genetics* **98,** 138-140.

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. *Human Molecular Genetics* **4,** 843-852.

Jansen, G., Groenen, P., Bachner, D., Jap, P. H. K., Coerwinkel, M., Oerlemans, F., Vandenbroek, W., Gohlsch, B., Pette, D., Plomp, J. J., Molenaar, P. C., Nederhoff, M. G. J., Vanechteld, C. J. A., Dekker, M., Berns, A., Hameister, H. and Wieringa, B. (1996). Abnormal myotonic dystrophy protein kinase levels produce only mild myopathy in mice. *Nature Genetics* **13,** 316-324.

Jansen, G., Mahadevan, M., Amemiya, C., Wormskamp, N., Segers, B., Hendriks, W., O'hoy, K., Baird, S., Sabourin, L., Lennon, G., Jap, P. L., Iles, D., Coerwinkel, M., Hofker, M., Carrano, A. V., de Jong, P. J., Korneluk, R. G. and Wieringa, B. (1992). Characterization of the myotonic dystrophy region predicts multiple protein isoform-encoding mRNAs. *Nature Genetics* **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)n repeat variation and selection against extreme expansion in sperm. *American Journal Of Human Genetics* **54,** 575-585.

Jaworski, A., Rosche, W. A., Gellibolian, R., Kang, S., Shimizu, M., Bowater, R., Sinden, R. R. and Wells, R. D. (1995). Mismatch repair in *Escherichia coli* enhances instability of $(CTG)_n$ triplet repeats from human hereditary diseases. *Proceedings Of the National Academy Of Sciences Of the United States Of America* **92,** 11019-11023.

Jodice, C., Malaspina, P., Persichetti, F., Novelletto, A., Spadaro, M., Giunti, P., Morocutti, C., Terrenato, L., Harding, A. E. and Frontali, M. (1994). Effect Of Trinucleotide Repeat Length and Parental Sex On Phenotypic Variation In Spinocerebellar Ataxia-I. *American Journal Of Human Genetics* **54,** 959-965.

Johnson, R. E., Kovvali, G. K., Guzder, S. N., Amin, N. S. and Holm, C. (1996). Evidence for involvement of yeast proliferating cell nuclear antigen in DNA mismatch repair. *Journal of Biological Chemistry* **271,** 27987-27990.

Joseph, J. T., Richards, C. S., Anthony, D. C., Upton, M., PerezAtayde, A. R. and Greenstein, P. (1997). Congenital myotonic dystrophy pathology and somatic mosaicism. *Neurology* **49,** 1457-1460.

Kang, S. M., Ohshima, K., Shimizu, M., Amirhaeri, S. and Wells, R. D. (1995). Pausing of DNA synthesis in vitro at specific loci in CTG and CGG triplet repeats from human hereditary disease genes. *Journal Of Biological Chemistry* **270,** 27014-27021.

Kato, T., Yatagai, F., Glickman, B. W., Tachibana, A. and Ikenaga, M. (1998). Specificity of mutaions in th PMS2-deficient human tumour cell line HEC-1-A. *Mutation Research* **422,** 279-283.

Kawaguchi, Y., Okamoto, T., Taniwaki, M., Aizawa, M., Inoue, M., Katayama, S., Kawakami, H., Nakamura, S., Nishimura, M., Akiguchi, I., Kimura, J., Narumiya, S. and

Kakizuka, A. (1994). CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nature Genetics* **8,** 221-228.

Kaytor, M. B., E.N. Duvick, L.A. Zoghbi, H.Y. Orr, H.T. (1997). Increased trinucleotide repeat instability with advanced maternal age. *Human Molecular Genetics* **6,** 2135-2139.

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. *Nature Genetics* **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. *Nature Genetics* **16,** 402-406.

Knight, S. J. L., Flannery, A. V., Hirst, M. C., Campbell, L., Christodoulou, Z., Phelps, S. R., Pointon, J., Middleton-Price, H., Barnicoat, A., Pembrey, M. E., Holland, J., Oostra, B. A., Bobrow, M. and Davies, K. E. (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. *Journal of Theoretical Biology* **192,** 505-14.

Kokoska, R. J., Stefanovic, L., Buermeyer, A. B., Liskay, R. M. and Petes, T. D. (1999). A mutation of the yeast gene encoding PCNA destabilzes both microsatellite and minisatellite DNA sequences. *Genetics* **151,** 511-19.

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). *Nature Genetics* **21,** 379-384.

Kovton, I. V., Therneau, T. M. and McMurray, C. T. (2000). Gender of the embryo contributes to CAG instability transgenic mice containing a Huntington's disease gene. *Human Molecular Genetics* **9,** 2767-2775.

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.

Krentz, A. J., Coles, N. H., Williams, A. C. and Nattrass, M. (1990). Abnormal Regulation Of Intermediary Metabolism After Oral Glucose Ingestion In Myotonic Dystrophy. *Metabolism-Clinical and Experimental* **39,** 938-942.

La Spada, A. R., Peterson, K. R., Meadows, S. A., McClain, M. E., G, J., Chmelar, R., Haugen, H. A., Chen, K., Singer, M. J., Moore, D., Trask, B. J., Fischbeck, K. H., Clegg, C. H. and McKnight, G. S. (1998). Androgen receptor YAC transgenic mice carrying CAG 45 alleles show trinucleotide repeat instability. *Human Molecular Genetics* **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.

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

Lavedan, C., Hofmannradvanyi, H., Rabes, J. P., Roume, J. and Junien, C. (1993b). Different Sex-Dependent Constraints In CTG Length Variation As Explanation For Congenital Myotonic Dystrophy. *Lancet* **341,** 237-237.

Leadon, S. A. and Avrutskaya, A. V. (1997). Differential involvement of the human mismatch repair proteins, hMLH1 and hMSH2, in transription coupled repair. *Cancer Research* **57,** 3784-3791.

Lia, A.-S., Seznec, H., Hofmann-Radvanyi, H., Radvanyi, F., 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. *Human Molecular Genetics* **7,** 1285-1291.

Liu, B. P., R. Papadopoulos, N, Nicolaides, NC, Lynch, H T. Watson, P. Jass, JR. Dunlop, M. Wyllie, A. Peltomaki, P. de la Chapelle, A. Hamilton, SR. Vogelstein, B. Kinzler, KW. (1996). Analysis of mismatch repair genes in heriditary non-polyposis colorectal cancer patients. *Nature Medicine* **2,** 169-174.

Lopez de Munain, A., Cobo, A. M., Pozo, J. J., Navarette, D., Martorell, L., Palau, F., Emparanza, J. I. and M, B. (1995). Influence of the sex of the transmitting grandparent in congenital myotonic dystrophy. *Journal of Medical Genetics* **32,** 689-691.

Lorenzetti, D., Watase, K., Xu, B., Matzuk, M. M., Orr, H. and Zoghbi, H. Y. (2000). Repeat instability and motor incoordination in mice with a targeted expanded CAG repeat in the Sca 1 locus. *Human Molecular Genetics* **9,** 779-785.

Lynch, H., Smyrk, T. and Lynch, J. (1997). An update of HNPCC (Lynch syndrome). *Cancer Genetics and Cytogentics* **93,** 84 - 99.

Lynch, H. T. and Smyrk, T. (1996). Hereditary Nonpolyposis Corectal Cancer. *Cancer* **78,** 1149-1163.

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

Mahadevan, M. S., Foitzik, M. A., Surh, L. C. and Korneluk, R. G. (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. *Nature Genetics* **15,** 197-200.

168

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 expnded CUG repeat. *Science* **289,** 1769-1772.

Manley, K., Shirley, T. L., Flaherty, L. and Messer, A. (1999). Msh2 deficiency prevents *in vivo* somatic intability of the CAG repeat in Huntington disease transgenic mice. *Nature Genetics* **23,** 471-473.

Martorell, L., Monckton, D. G., Gamez, J. and Baiget, M. (2000). Complex patterns of germline instability and somatic mosiacism in myotonic dystrophy type 1. *European Journal of Human Genetics* **8,** 423-430.

Martorell, L., Monckton, D. G., Gamez, J., Johnson, K. J., Gich, I., Lopez de Munain, A. and Baiget, M. (1998). Progression of somatic CTG repeat length heterogeneity in the blood cells of myotonic dystrophy patients. *Human Molecular Genetics* **7,** 307-312.

Massari, A., Gennarelli, M., Menegazzo, E., Pizzuti, A., Silani, V., Mastrogiacomo, I., Pagani, E., Angelini, C., Scarlato, G., Novelli, G. and Dallapiccola, B. (1995). Postzygotic instability of the myotonic dystrophy repeat supported by larger expansions in muscle and reduced amplifications in sperm. *Journal Of Neurology* **242,** 379-383.

McCarthy, G. and Leblond, C. (1988). Radioautographic evidence for slow astrocyte turnover and modest oligodendrocyte production in the corpus callusum of adult mice infused with 3H-thymidine. *Journal of Comparative Neurology* **271,** 589-603.

Mellon, I. and Champe, G. (1996). Products of DNA mismatch repair genes mutS and mutL are required for transcripton-coupled nucleotide-excision repair of the lactose operon in *Escherichia coli. Proceedings Of the National Academy Of Sciences Of the United States Of America* **93,** 1292-1297.

Messier, B. and Leblond, C. P. (1960). Cell proliferation and migration as revealed by autoradiography after injection of thymidine H3 into male rats and mice. *American Journal of Anatomy* **106,** 247-265.

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 Research* **27,** 3534-3542.

Miller, J. W., Urbinati, C. R., Teng-umnuay, P., Stenburg, 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. *European Molecular Biology Organisation* **19,** 4439-4448.

Modrich, P. and Lahue, R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annual Review of Biochemistry* **65,** 101-133.

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. *Nature Genetics* **15,** 193-196.

Monckton, D. G., Wong, L. J. C., Ashizawa, T. and Caskey, C. T. (1995). Somatic mosaicism, germline expansions, germline reversions and intergenerational reductions in myotonic dystrophy males small pool pcr analyses. *Human Molecular Genetics* **4,** 1-8.

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. *Human Molecular Genetics* **,** 2313-2320.

Moxley, R. T., Kingston, W. J., Griggs, R. C. and Livingston, J. N. (1987). Lack Of Rapid Enhancement Of Insulin Action After Oral Glucose Challenge In Myotonic Dystrophy. *Diabetes* **36,** 693-701.

Nagafuchi, S., Yanagisawa, H., Ohsaki, E., Shirayama, T., Tadokoro, K., Inoue, T. and Yamada, M. (1994). Structure and expression of the gene responsible for the triplet repeat disorder, dentatorubral and pallidoluysian atrophy (DRPLA). *Nature Genetics* **8,** 177-182.

Narayanan, L. F., J.A. Baker, S.M. Liskay, R.M. Glazer P.M. (1997). Elevated levels of mutation in multiple tissues of mice deficient in the DNA mismatch repair gene *Pms2*. *Genetics* **94**, 3122-3127.

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. *Human Molecular Genetics* **3**, 45-51.

Ohshima, K., Kang, S., Larson, J. E. and Wells, R. D. (1996). Cloning, characterization, and properties of seven triplet repeat DNA sequences. *Journal of Biological Chemistry* **271**, 16773-16783.

Ohzeki, S., Tachibana, A., Tatsumi, T. and Kato, T. (1997). Spectra of spontaneous mutations at the *hprt* locus in colorectal carcinoma cell lin defective in mismatch repair. *Carcinogenesis* **18**, 1127-1133.

Orr, H. T., Chung, M. Y., Banfi, S., Kwiatkowski, T. J., Servadio, A., Beaudet, A. L., McCall, A. E., Duvick, L. A., Ranum, L. P. W. and Zoghbi, H. Y. (1993). Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type-1. *Nature Genetics* **4**, 221-226.

Otten, A. D. and Tapscott, S. J. (1995). Triplet repeat expansion in myotonic dystrophy alters the adjacent chromatin structure. *Proceedings Of the National Academy Of Sciences Of the United States Of America* **92**, 5465-5469.

Paraf, F., Jothy, S. and Van Meir, E. (1997). Brain tumour-polyposis syndrome: two genetic diseases. *Journal of Clinical Oncology* **15**, 2744-2758.

Parniewski, P., Jaworski, A., Wells, R. and Bowater, R. (2000). Length of CAG◊CTG repeats determines the influence of mismatch repair on genetic instability. *Journal of Molecular Biology* **299**, 865-874.

Parsons, M. A., Sinden, R. R. and Izban, M. G. (1998). Transcriptional properties of RNA polymerase II within triplet repeat-containing DNA from the human myotonic dystrophy and fragile X loci. *Journal of Biolological Chemistry* **273**, 26998-7008.

Pearson, C. E., Wang, Y. H., Griffith, J. D. and Sinden, R. R. (1998). Structural analysis of slipped strand DNA (S DNA) formed in $(CTG)_{(n)}/(CAG)_{(n)}$ repeats from the myotonic dystrophy locus. *Nucleic Acids Research* **26**, 816-823.

Peltomaki, P. and de la Chappelle, A. (1997). Mutations predisposing to hereditary non-polyposis corectal cancer. *Advances in Cancer Research* **71**, 93-119.

Peltomaki, P. and Vasen, H. (1997). Mutations predisposing to herditary nonpolyposis correctal cancer: database and results of a collaborative study. The International Collaborative Group on Heriditary Nonpolyposis Colorectal Cancer. *Gastroenterology* **113**, 1146-1158.

Penrose, L. S. (1948). The problem of anticipation in pedigrees of dystrophia myotonica. *Annals of Eugenics* **14**, 125-132.

Phillips, 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.

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., Shibata, D., Bradley, A. and Liskay, R. M. (1998). Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair. *Nature Genetics* **18**, 276-279.

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

Qin, X., Liu, L. and Gerson, S. (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.

Rayssiguier, C., Thaller, D. S. and Radman, M. (1989). The barrier to recombination between *Escherichia coli* and *Salmonella typhimurium* is disrupted in mismatch repair mutants. *Nature* **342,** 396-401.

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

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

Roses, A. D., Taylor, H., Matthews, P., Schwartzbach, C., Kozataylor, P., Potter, T., Stajick, J., Gaskell, P., Friedman, A., Speer, M., Pericakvance, M., Xu, P. T. and Gilbert, J. (1996). Difference in the expression of myotonin protein kinase messenger RNA in adult and congenital myotonic muscular dystrophy. *Neurology* **46,** 68002.

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

Sadler, T. W. (1985). Medical Embryology: Williams and Wilkins.

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

Sarkar, P. S., Appukuttan, B., Han, J., Ito, Y., Ai, C., Tsai, W., Chai, Y., Stout, J. T. and Reddy, S. (2000a). Heterozygous loss of Six5 in mice is sufficient to cause ocular cataracts. *Nature Genetics* **25,** 110-114.

Sarkar, P. S., Han, J. and Reddy, S. (2000b). DM 1 Workshop. In *American Society of Human Genetics*. Philadelphia.

Sasagawa, N., Takahashi, N., Suzuki, K. and Ishiura, S. (1999). An expanded CTG trinucleotide repeat causes trans RNA interference: a new hypothesis for the pathogenesis of myotonic dystrophy. *Biochemica Biophysica Research Communication* **264,** 76-80.

Sato, T., Oyake, M., Nakamura, K., Nakao, K., Fukusima, Y., Onodera, O., Igarashi, S., Takano, H., Kikugawa, K., Ishida, Y., Shimohata, T., Koide, R., Ikeuchi, T., Tanaka, H., Futamura, N., Matsumura, R., Takayanagi, T., Tanka, F., Sobue, G., Komure, O., Takahashi, M., Sano, A., Ichikawa, Y., Goto, J., Kanazwa, I., Katsuki, M. and Tsuji, S. (1999). Transgenic mice harboring a full length human mutant DRPLA gene exhibit age-dependant intergenerational and somatic instabilities of CAG repeats comparable with those in DRPLA patients. *Human Molecular Genetics* **8,** 99-106.

Schumacher, S., Fuchs, R. P. P. and Bichara, M. (1997). Two distinct models account for long and short deletions within sequence repats in *Escherichia coli*. *Journal of Bacteriology* **279,** 1101-1110.

Schweitzer, J. K. and Livingston, D. M. (1997). Destabilization of CAG trinucleotide repeat tracts by mismatch repair mutations in yeast. *Human Molecular Genetics* **6,** 349-355.

Selva, E., New, L., Crouse, G. and Lahue, R. (1995). Mismatch correction acts as a barrier to homologous recombination in *Sacchromyces cervisiae*. *Genetics* **139,** 1175-1188.

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. *Human Molecular Genetics* **9,** 1185-1194.

Shaw, D. J., McCurrach, M., Rundle, S. A., Harley, H. G., Crow, S. R., Sohn, R., Thirion, J. P., Hamshere, M. G., Buckler, A. J., Harper, P. S., Housman, D. E. and Brook, J. D. (1993). Genomic organization and transcriptional units at the myotonic dystrophy locus. *Genomics* **18,** 673-679.

Shaw, J. A., Walsh, T., Chappell, S. A., Carey, N., Johnson, K. and Walker, R. A. (1996). Microsatellite instability in early sporadic breast cancer. *British Journal Of Cancer* **73**, 1393-1397.

Shelbourne, P., Killeen, N., Hevner, R., Johnston, H. and Tecott, L. (1999). A Huntington's disease CAG at the murine Hdh locus is unstable and associated with behavioural abnormalities in mice. *Human Molecular Genetics* **8**, 763-774.

Sinden, R. R. (1999). Biological implications of the DNA structures associated with disease-causing triplet repeats. *American Journal of Human Genetics* **64**, 346-53.

Steinbach, P., Glaser, D., Vogel, W., Wolf, M. and Schwemmle, S. (1998). The DMPK gene of severely affected myotonic dystrophy patients is hypermethylated proximal to the large expanded CTG repeat. *American Journal of Human Genetics* **62**, 278-283.

Steinert, H. (1909). Uber das klinische und anatomische bild des muskelschwunds der myotoniker. *Dtsch. Z. Nervenheilkd* **37**, 58-104.

Stevanin, G., Cancel, G., Durr, A., Chneiweiss, H., Dubourg, O., Weissenbach, J., Cann, H. M., Agid, Y. and Brice, A. (1995). The gene for spinal cerebellar ataxia 3 (SCA 3) is located in a region of ~3cM on chromosome 14q24.3-q32.2. *American Journal of Human Genetics* **56**, 193-201.

Tachi, N., Kozuka, N., Ohya, K., Chiba, S. and Kikuchi, K. (1995a). Expression of myotonic dystrophy protein kinase in biopsied muscles. *Journal Of the Neurological Sciences* **132**, 61-64.

Tachi, N., Ohya, K., Chiba, S. and Sato, T. (1993). Unstable DNA In a Patient With a Severe Form Of Congenital Myotonic Dystrophy. *Journal Of the Neurological Sciences* **119**, 180-182.

Tachi, N., Ohya, K., Chiba, S., Sato, T. and Kikuchi, K. (1995b). Minimal somatic instability of CTG repeat in congenital myotonic dystrophy. *Pediatric Neurology* **12**, 81-83.

Tachi, N., Ohya, K., Yamagata, H., Miki, T., Kikuchi, K. and Chiba, S. (1997). Haplotype analysis of congenital myotonic dystrophy patients from asymptomatic DM father. *Pediatric Neurology* **16,** 315-318.

Taketo, M., Schroeder, A., Mobraaten, L., Gunning, K., Hanten, G., Fox, R., Roderick, T., Stewart, C., Lilly, F., Hansen, C. and Overbeek, P. (1991). FVB/N: An inbred mouse strain preferable for transgenic analyses. *Proc. Natl. Acad. Sci. USA* **88,** 2065-2069.

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. *Journal Of Cell Biology* **128,** 995-1002.

Thornton, C. A., Wymer, J. P., Simmons, Z., McClain, C. and Moxley, R. T. r. (1997). Expansion of the myotonic dystrophy CTG repeat reduces expression of the flanking DMAHP gene. *Nature Genetics* **16,** 407-9.

Tian, B., White, R. J., Tianbang, X., 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-dependant protein kinase PKR. *RNA* **6,** 79-87.

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,** 4517-25.

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. *Nature Genetics* **1,** 192-195.

Tsuji, S. (1997). Molecular gentics of triplet repeat diseases: Unstable expansion of triplet repeats as a new mechanism for neurodegenerative diseases. *Internal Medicine* **36,** 3-8.

Umar, A., Boyer, J. C. and Kunkel, T. A. (1994a). DNA loop repair by human cell extracts. **Science,** 814-816.

Umar, A., Buermeyer, A. B., Simon, J. A., Thomas, D. C. and Clark, A. B. (1994b). Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell* **87**.

Usuki, K. and Ishiura, S. (1998). Expanded CTG repeats in myotonin protein kinase increase susceptibility to oxiditve stress. *Clinical Neuroscience* **9**, 2291-2296.

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

Wang, J., Morrone, A., Pegoraro, E., Angelini, C., Zammarchi, E., Marconi, G. and Hoffman, E. P. (1995). Myotonic dystrophy: is this an RNA expansion disorder? *American Journal of Human Genetics* **57 Supplement**, A1469.

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.

Watkins, W. S., Bamshad, M. and Jorde, L. B. (1995). Population genetics of trinucleotide repeat polymorphisms. *Human Molecular Genetics* **4**, 1485-1491.

Wells, R. D. (1996). Molecular basis of genetic instability of triplet repeats. *Journal of Biological Chemistry* **271**, 2875-2878.

Wheeler, V. C., Auerbacg, W., White, J. K., Srinidhi, J., Auerbach, A., Ryan, A., Duyao, M. P., Vrbanac, V., Weaver, M., Gusella, J. F., L, J. A. and E, M. M. (1999). Length-dependent gametic CAG repeat instabilty in the Huntington's disease knock in mouse. *Human Molecular Genetics* **8**, 115-122.

Wieringa, B. (1994). Commentary - Myotonic Dystrophy Reviewed - Back to the Future. *Human Molecular Genetics* **3,** 1-7.

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

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

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

Yamagata, H., Miki, T., Nakagawa, M., Johnson, K., Deka, R. and Ogihara, T. (1996). Association of CTG repeats and the 1 kb alu insertion deletion polymorphism at the myotonin protein kinase gene in the japanese population suggests a common eurasian origin of the myotonic dystrophy mutation. *Human Genetics* **97,** 145-147.

Yao, X., Buermeyer, A. B., Narayanan, L., Tran, D., Baker, S. M., Prolla, T. A., Glazer, P. M., Liskay, R. M. and Arnheim, N. (1999). Different mutator phenotypes in Mlh1- versus Pms2-deficient mice. *Proceedings Of the National Academy Of Sciences Of the United States Of America* **96,** 6850-5.

Yu, A., Dill, J., Wirth, S. S., Huang, G., Lee, V. H., Haworth, I. S. and Mitas, M. (1995). The trinucleotide repeat sequence d(GTC)15 adopts a hairpin conformation. *Nucleic Acids Research* **23,** 2706-2714.

Zerylnick, C., Torroni, A., Sherman, S. L. and Warren, S. T. (1995). Normal variation at the myotonic dystrophy locus in global human populations. *American Journal of Human Genetics* **56,** 123-130.

Zhuchenko, O. B., J. Bonnen, P. Ashizawa, T. Stockton, D.W., Amos, C. Dobyns, Subramony, S.H. Zoghbi, H.Y. Lee, C.C. (1997). Autosomal dominant cerebullar ataxia (SCA 6) associated with small polyglutamine expansions in the a_{1A} -voltage dependant calcium channel. *Nature Genetics* **15,** 62-69.

Zuhlke, C., Riess, O., Bockel, B., Lange, H. and Thies, U. (1993). Mitotic Stability and Meiotic Variability Of the (CAG)(N) Repeat In the Huntington Disease Gene. *Human Molecular Genetics* **2,** 2063-2067.