TELOMERE STABILITY IN CANCER CELLS AND THE ROLE OF DNA MISMATCH REPAIR

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by
Aarón Méndez Bermúdez
Department of Genetics

September 2007

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

Contents

Chapter 2 Materials and Methods

Chapter 3 Telomere Mutation Analysis in Mismatch Repair Deficient Cell Lines

Chapter 4 Analysis of Telomere Stability in Colorectal Tissue Using a Single Molecule Approach

Chapter 5 Downregulation of *MSH2* Expression by shRNA

Chapter 7 Telomere Instability in Liposarcomas with Unknown Telomere Maintenance Mechanism

TELOMERE STABILITY IN CANCER CELLS AND THE ROLE OF DNA MISMATCH REPAIR

Aaron Mendez Bermudez

Abstract

Human telomeres are essential for maintaining chromosome stability and are composed of TTAGGG repeat arrays with interspersed variant repeats at the proximal end. This work has focused on understanding the role of DNA mismatch repair on telomere stability by mapping the interspersion pattern of variant telomere repeats in search of mutations. Telomere instability was analysed in clones of mismatch repair (MMR) defective cell lines and colorectal carcinomas with *MSH2* mutations. The combined mutation frequency found in four MMR-defective cell lines (3.6% per telomere per genome) and two colorectal tumours lacking the *MSH2* gene (1.2% per telomere per genome) was significantly different compared to MMR proficient cell lines. The mutations resulted in losses, gains and telomere repeat-type changes suggesting that mutations originated by intra-allelic mechanisms. In addition, downregulation of the *MSH2* gene was carried out in a normal fibroblast cell line using RNA interference. Downregulation of *MSH2* caused microsatellite and telomere instability with a possible effect on telomere length. These results indicate that the MMR pathway is involved in maintaining the stability of telomeres.

One variant telomere repeat found in humans is the CTAGGG repeat. When multiple tandem copies of CTAGGG are present in a telomere, it is highly unstable in the male germ-line with a mutation rate as high as 0.83 per sperm. In this work, the analysis of instability of CTAGGG-containing telomeres was extended in the germ-line and conducted in somatic tissue. The results indicated that the presence of more than five CTAGGG repeats cause instability in that telomere, with a mutation frequency of 3.8% per telomere per genome in somatic cells.

Finally, the basis of the telomere maintenance mechanism (TMM) was investigated in a panel of liposarcomas where around 50% are reported to lack a TMM. By mapping telomeres from liposarcomas with unknown TMM, it was possible to identify complex mutations that arose, most probably, by recombination-like processes similar to that seen in the alternative lengthening of telomeres pathway (ALT). However, these tumours lack some of the markers for ALT such as ALT-associated PML bodies.

Acknowledgements

Firstly I would like to thank Dr Nicola Royle for being an excellent supervisor, for her help and guidance during my PhD. Thanks also to Jennie and Jenny for much appreciated help. Thanks also go to all lab members past and present.

Finally, I would like to thank the Mexican Council of Science and Technology (CONACyT) for financial support.

Abbreviations

bp	base pair
cDNA	complementary DNA
DNA	deoxyribonucleic acid
dsRNA	double-stranded RNA
kDa	kilo Daltons
Mb	mega base
MEM	minimal essential medium
mRNA	messenger RNA
μ	mutation rate
μF	micro Faraday
μl, ml, l	micro-, mili-, litre
μM, mM, M	micro-, mili-, molar
nt	nucleotide
OD	optical density
PCR	polymerase chain reaction
PD	population doubling
PBS	phosphate buffered saline
pg, ng, μg	pico-, nano-, microgram
pmols	picomols
RNA	ribonucleic acid
rpm	revolutions per minute
V	volts

Publications

Some of this work has been published.

Jeyapalan J, Mendez-Bermudez A, Zaffaroni N, Dubrova Y and Royle NJ. Evidence for Alternative Lengthening of Telomeres in liposarcomas in the absence of ALT associated PML bodies. *International Journal of Cancer* (Paper accepted).

Stagno D'Alcontres M, Mendez-Bermudez A, Foxon JL, Royle NJ and Salomoni P. (2007) Lack of TRF2 in ALT cells causes PML-dependent p53 activation and loss of telomeric DNA. *Journal of Cell Biology* **173**, 855-867.

Chapter 1

Introduction

1.1 Structure of telomeres

Telomeres are the DNA-protein structures located at the ends of eukaryotic chromosomes. Telomere research was initiated by the study of chromosome stability in *Drosophila melanogaster* by Herman Muller and in *Zea mays* by Barbara McClintock (Muller, 1938; McClintock, 1941). Muller observed different types of chromosome rearrangements after *Drosophila* cells were exposed to X-ray radiation and he deduced that the terminal DNA of linear chromosomes must have a specialised structure that protects the chromosome ends (Muller, 1938). McClintock demonstrated in maize that broken chromosomes are unstable, undergoing aberrant recombination and end-to-end fusions. Furthermore, McClintock showed that in some situations, broken chromosomes can be healed, presumably by the addition of telomeres, resulting in normal chromosome behaviour (McClintock, 1941). These observations were the first pieces of evidence that suggested chromosome ends are different from double strand breaks and are important for chromosome stability.

1.1.1 Telomeres are formed by tandem repeat arrays

The structure of telomeres was first characterised in *Tetrahymena* (Blackburn and Gall, 1978) where the TTGGGG sequence was found organised in tandem arrays at the end of chromosomal DNA (Table 1.1). As for *Tetrahymena*, sequence analysis revealed similar tandem repeat organisation of the DNA at the end of linear molecules in the ciliate *Oxytricha*, with a consensus sequence of T_4G_4 (Klobutcher *et al.*, 1981). Studies in different vertebrates showed that telomere sequence and organisation are highly conserved (Meyne *et al.*, 1989) with the G-rich DNA strand orientated 5' → 3' towards the end of the chromosomes and forming a 3' overhang.

Organism	Telomere sequence	Telomere length	Reference
Tetrahymena (protozoa)	TTGGGG	Ma: 250 – 350 bp Mi: 2 – 3.4 kb	Blackburn and Gall, 1978; Kirk and Blackburn, 1995
Oxytricha (protozoa)	TTTTGGGG	Ma: 28 or 20 bp Mi: 3 – 20 kb	Klobutcher *et al.*, 1981
Saccharomyces (fungi)	$G_{2-3}(TG)_{1-6}$	200 – 300 bp	Shampay *et al.*, 1984
Schizosaccharomyces (fungi)	GGTTACA[1]	~300 bp	Hiraoka *et al.*, 1998
Caenorhabditis (invertebrates)	TTAGGC	4 - 9 kb	Wicky *et al.*, 1996
Drosophila (invertebrates)	Retrotransposon sequences	Variable size[2]	Sheen and Levis, 1994
Chlamydomonas (plants)	TTTTAGGG	300 – 350 bp	Petracek *et al.*, 1990
Arabidopsis (plants)	TTTAGGG	2 – 9 kb	Richards and Ausubel, 1998
Mouse	TTAGGG	> 30 kb[3]	Starling *et al.*, 1990
Humans	TTAGGG	2 – 20 kb	Moyzis *et al.*, 1988; Counter *et al.*, 1992

Table 1.1 Telomere sequence and length of some of the most common model organisms. Budding and fission yeast telomeres are irregular and contain variant repeats such as the $T_{1-2}ACA_{0-1}C_{0-1}G_{1-6}$ sequence in *Schizosaccharomyces*[1]. *Drosophila* telomeres are formed by retrotransposon sequences of a few kb each and their telomeres show variable length[2]. The length of mouse telomeres vary greatly, however most laboratory mouse strains have telomeres longer than 30 kb[3]. Abbreviations are Ma: Macro-nucleus, Mi: Micro-nucleus.

1.1.2 Organisation of subtelomeric DNA

The basic unit of eukaryotic chromatin is the nucleosome. This structure is formed by an octameric core of histones (H2A, H2B, H3 and H4) and approximately 200 bp of DNA which is wrapped around the outside of the histone octamer. The nucleosomes are coiled and packaged forming two types of chromatin: the euchromatin, an open and relax chromatin which is associated with active genes and the heterochromatin which is condensed and is normally

associated with repressed transcription sites. It is believed that modifications in the N-terminal tail of histones (such as methylations, acetylations and phosphorylations) create conformational changes of nucleosomes than can change the structure of chromatin (reviewed in Khorasanizadeh, 2004).

Subtelomeres are low in gene density containing middle-repetitive elements called telomeric-associated sequences with DNA similar to heterochromatin. The subtelomeric regions in *S. cerevisiae* are some of the best characterised. Each comprises a ~25 kb region and is formed by two main elements: the Y' and the conserved core X element. The Y' elements are not found at all chromosome ends, however up to four of them can be found in a single subtelomeric region. They are organised in tandem arrays separated by short regions of telomeric repeats. There is no apparent function for the Y' in wild-type yeast strains, however, in telomerase deficient strains the Y' exhibit high levels of rearrangements as seen by an increase in copy number and distribution. These rearrangements enhance cell viability (further details are discussed below) (Lundblad and Blackburn, 1993). The other major component of budding yeast subtelomeres is the core X that is found in all chromosome ends. This element is 473 bp in length and it is found proximal to the Y' elements. The core X has binding sites for the yeast origin recognition complex (ORC) and the Abf1 transcription factor, thus this element could have functional roles at telomeres (Louis, 1995; Pryde and Louis, 1999).

Human subtelomeres, also known as proterminal regions, are formed by repetitive sequences, internal telomeric sequences and segments of duplicated DNA that vary from 1 kb to 200 kb. In zygotene of meiosis I, the formation of the synaptonemal complex required for the physical association of homologous chromosome is initiated at subtelomeres, thus these structures are regions with high levels of meiotic recombination (Royle, 1995; Riethman *et al.*, 2001; Brown *et al.*, 2005). Human subtelomeric regions are rich in transcripts, some of them containing zinc finger protein domains although for the majority of them it is unknown if they produce functional proteins (Riethman *et al.*, 2004). Thus, it has been hypothesised that subtelomeres are functional part of the chromosomes although much of their function is still unknown. Other components of subtelomeres, such as the internal $(TTAGGG)_n$-like repeats, are extremely variable and are thought to originate from subtelomeric recombination events (Mondello *et al.*, 2000).

1.2 Replication of telomeres

All eukaryote organisms maintain their genomes as linear DNA molecules in chromosomes. The replication of the full linear DNA molecule requires special mechanisms in order to maintain its integrity for two reasons. Firstly, DNA replication is semi-conservative, which means synthesis of a new strand requires a parental strand template. Since eukaryotic telomeres end in a 3' overhang, the 5' end strand cannot provide a template for the synthesis of the 3' overhang. Thus, creating an incompletely replicated chromosome end (Figure 1.1). Second, all DNA polymerases can only replicate DNA in the 5' to 3' direction and most cannot synthesis DNA *de novo*. The DNA polymerases δ (mainly active in the lagging strand) and ε (in the leading strand) uses a short sequence of RNA (8 to 12 nucleotides long) to initiate DNA synthesis. One strand is synthesised completely as a continuous process (leading strand synthesis) whereas the other strand is synthesised as discontinuous fragments (Okazaki fragments) that are subsequently ligated together (lagging strand synthesis). After synthesis has taken place, the RNA primer that is used to initiate the synthesis is removed leading to incomplete replication in the lagging strand and seen as a 3' overhang (Figure 1.1). However, since the lagging and the leading daughter strands contain 3' overhangs, it is proposed that in addition to the gap generated when the RNA primer is removed, the leading and probably the lagging strand are processed in order to either generate or enlarge the 3' overhang (Chai *et al.*, 2006; Sfeir *et al.*, 2005). Hence, in every cell cycle the end of the chromosomes will shorten due to the lack of complete replication and after a number of cell divisions without a mechanism to replicate telomeres fully, telomeres would become too short to fulfil their capping function so compromising chromosome integrity. Studies in human primary cells have shown that the rate of telomere loss per cell division is proportional to the size of the 3' overhang. For example cells with a 300 bp overhang lose 100 bp of sequence per cell division while cell with shorter overhangs of ~100 bp lose only 50 bp (Huffman *et al.*, 2000).

1.2.1 G-rich telomere overhang

As mentioned above, telomeres are composed of double-strand DNA followed by a single-strand overhang at the 3' end of the G-rich DNA strand. It is most likely that the generation of the 3' overhang is the result of incomplete replication and postreplicative processing (Figure 1.1). In organisms such as ciliates, the length and the terminal nucleotides of the G-overhangs are almost fixed. For example in *Tetrahymena*, the overhangs are either 14 - 15

or 20 - 21 nucleotides long. In addition the G-rich strand ends in TGGGGT-3' while the C-rich strand ends either in CAACCC-5' or CCAACC-5' (Jacob *et al.*, 2001). In contrast, the human G-rich overhangs contain variable terminal nucleotides. Nonetheless, it has been suggested that ~80% of the C-rich strand ends terminate in CCAATC-5' suggesting that a specific nuclease processing is acting in the C-rich strand after replication (Sfeir *et al.*, 2005).

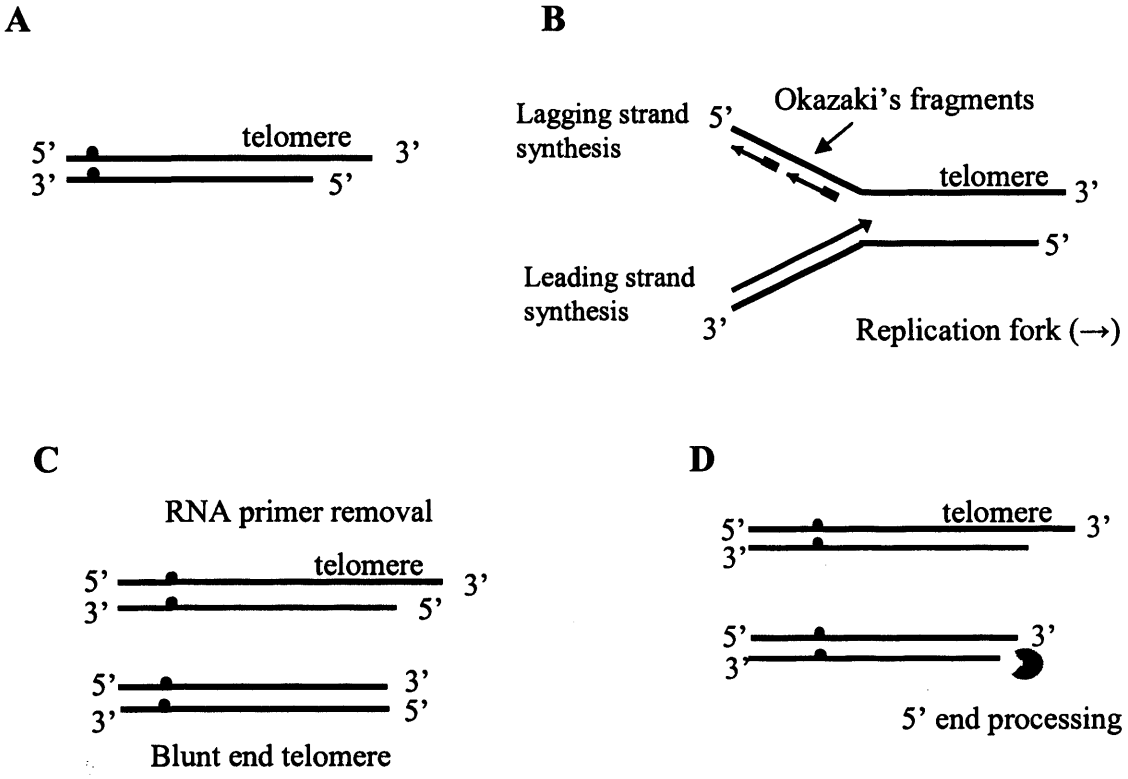


Figure 1.1 Diagram representing the replication process of telomeres. Eukaryotic chromosomes end in a 3' overhang (A). The lagging strand synthesis is discontinuous, being primed by short RNA oligonucleotides (B). At the end of replication, the RNA oligonucleotides are removed and fill by DNA polymerases. A gap is created at the end of the telomere of the lagging strand, while the leading strand synthesis produces blunt ends (C). Ultimately, nucleases process the 5' end to generate 3' overhangs (D). (→) direction of the replication fork.

The length of the G-overhangs in human cells is variable but normally no longer than 200 nt (Wright *et al.*, 1997; Chai *et al.*, 2006a) differing from some model organisms where the

length of the overhangs is fixed and short. In *S. cerevisiae*, for example, the G-overhangs are always 12 to 14 nt generated most probably by the removal of the RNA primer in lagging strand synthesis (Figure 1.1). Another difference between the telomere overhangs of yeast and humans is that in human cells the size of the G-overhangs varies between the leading and the lagging daughter telomeres with longer overhangs in the latter. In Chai *et al.* (2006) the authors allowed fibroblast cells to incorporate 5-bromodeoxyuridine (analogue thymidine) for one cell division. Therefore, the leading-strand daughter telomere is heavier than the lagging-daughter strand and can be separate by density gradient. Subsequently, both daughter-telomere overhangs were measured in two normal human cell lines (BJ and IMR90 fibroblasts) finding in both cell lines that the leading daughter strand overhang was shorter and about half the length of the overhang on the lagging daughter strand (60 nt versus 105 nt in BJ cells and 65 nt versus 110 nt in IMR90 cells). In telomerase positive cell lines, the length of the G-overhang is similar between the leading and the lagging daughter strand, suggesting telomerase adds telomere repeats preferentially at the 3' overhang of the leading daughter strand and then keeps constant the length at both daughter strands (Chai *et al.*, 2006a). It has also been observed that the recombination complex MRX (Mre11/Rad50/Xrs2) in *S. cerevisiae* and the MRN complex (Mre11/Rad50/Nbs1) in humans are involved in the generation of G-overhangs. In the yeast *S. cerevisiae*, the size of the G-overhangs range from 12 to 14 nt. However, deletion of the *MRE11* gene reduced the size of the G-overhangs to less than 8 nt linking the MRX complex with the generation and maintenance of G-overhangs (Larrivee *et al.*, 2004). Furthermore, the MRX complex was recruited to telomeres in late S phase, where it interacted with the telomere single-strand binding protein, Cdc13p and telomerase. Thus, this complex may regulate the elongation of the G-overhang in S phase through a telomerase mediated process (Takata *et al.*, 2005). Similar to yeast, in telomerase positive human cells, downregulation of the MRN complex by RNAi induced length reduction of the G-overhang. However, this effect was not observed in telomerase negative cells. It has been proposed that the MRN complex recruits telomerase to the end of the telomere, facilitating the generation of G-overhangs (Chai *et al.*, 2006b).

1.3 Proteins associated with telomeres and their role in telomere length regulation

Several proteins are associated with telomeres and regulate their length. Among these proteins, some bind double or single-stranded telomeric DNA directly. Double-stranded DNA

binding proteins include Rap1 (Repressor activator protein 1; found in *Saccharomyces cerevisiae*), Taz1p (found in *Schizosaccharomyces pombe*), Trf1 and Trf2 (TTAGGG Repeat binding factor 1 and 2; in mammals). Single-strand binding proteins include Cdc13p in the yeast *S. cerevisiae*, which binds at the 3' overhang of the telomeres. In the same way, Pot1 (Protection of telomeres 1) binds the single-stranded telomeric DNA in *S. pombe* and humans. Apart from proteins that bind telomeric repeats directly, there are other proteins that bind and link together Trf1, Trf2 and Pot1. In mammals, theses proteins include Tin2, Rap1 and Tpp1 and all together form a complex referred to as the telosome (Liu *et al.*, 2004) or shelterin (de Lange, 2005).

1.3.1 Description of the proteins found in the shelterin complex

In humans, shelterin is composed of 6 proteins (Trf1, Trf2, Pot1, Tin2, Tpp1 and Rap1; Figure 1.2) and can be isolated as single complex (Liu *et al.*, 2004; Ye *et al.*, 2004). Numerous other proteins are found at telomeres; however shelterin is characterised for being present throughout all phases of the cell cycle with functions limited to telomere homeostasis.

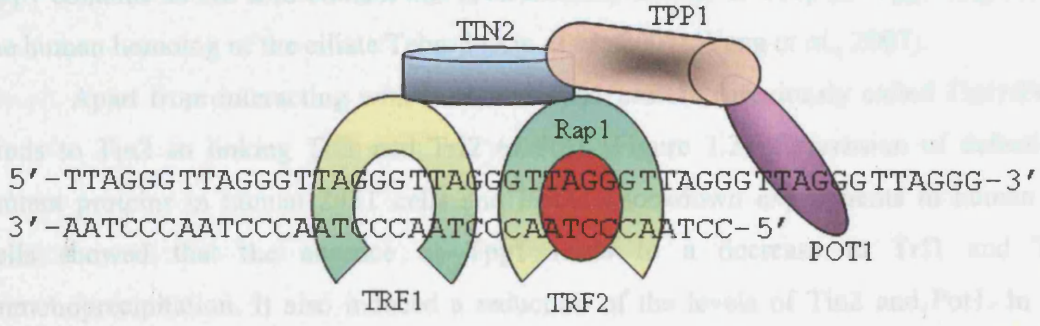


Figure 1.2 Diagram showing the shelterin members and their main interactions. Trf1, Trf2 and Pot1 proteins bind DNA directly, while Tin2 and Tpp1 probably link and stabilise the whole shelterin complex. Rap1 is recruited to telomeres through Trf2. Diagram modified from de Lange, 2005.

The members of the shelterin complex, Trf1 and Trf2 contain a Myb-type helix-turn-helix binding domain which can bind to the sequence motif 3'- YTAGGGTTR -5' (position Y can bind T or G while position R can bind A or C nucleotide; Bianchi *et al.*, 1999; Hanaoka *et*

al., 2005). Both proteins form stable homodimers that bind double-stranded telomeric repeats with the same recognition domain, although Trf1 has a greater DNA binding affinity (Hanaoka *et al.*, 2005). In contrast, Pot1 interacts with single-stranded telomeric DNA through 2 oligonucleotide/oligosaccharide binding fold (OB fold) domains; these domains have a minimal DNA binding site of 5'-TAGGGTTAG -3' (Liu *et al.*, 2004; Loayza *et al.*, 2004). Another protein that is found in the human shelterin complex is the yeast homolog of Rap1. Human Rap1 does not bind telomeric DNA instead it is recruited to telomeres through interactions with Trf2. The recruitment of Rap1 to telomeres *via* Trf2 interactions was demonstrated by expressing mutant forms of Rap1 that lacked its Trf2 interacting domain and by depletion of Trf2 from telomeres. In either case, Rap1 failed to localise to telomeres and its expression level declined (Celli and de Lange, 2005).

In the ciliate *Oxytricha nova* a heterodimer formed by Tebp-α (homolog of the mammalian Pot1) and Tebp-β was found to associate with single-stranded telomeric DNA, however only Tebp-α binds telomeric DNA sequences. Tebp-β interacts with Tebp-α and stabilises the complex (Gray *et al.*, 1991). In human cell extracts, co-immunoprecipitation studies have revealed the interaction of Pot1 with Tpp1 protein (Liu *et al.*, 2004). Furthermore Tpp1 contains an OB fold domain that is structurally similar to Tebp-β, suggesting that Tpp1 is the human homolog of the ciliate Tebp-β (Xin *et al.*, 2007; Wang *et al.*, 2007).

Apart from interacting with Pot1, the Tpp1 protein (previously called Tint1/Ptop/Pip1) binds to Tin2 so linking Trf1 and Trf2 to Pot1 (Figure 1.2). Expression of defective Tpp1 mutant proteins in human 293T cells and RNAi knockdown experiments in human HT1080 cells showed that the absence of Tpp1 leads to a decrease in Trf1 and Trf2 co-immunoprecipitation. It also induced a reduction of the levels of Tin2 and Pot1. In addition, depletion of Tin2 inhibited the formation of the shelterin complex. Therefore it is proposed that the complex formed by Tin2-Tpp1 is a key player in promoting and stabilising the shelterin complex (O'Connor *et al.*, 2006).

Shelterin proteins are found at telomeres at all stages of the cell cycle. However, there are other proteins that interact to telomeres from time to time and have important functions in establishing and maintaining the shelterin complex. Tankyrase 1 and 2 are among them. These proteins are poly(ADP-ribose) polymerases that inhibit the binding of Trf1 to telomeres by the addition of ADP-ribosylation polymers. Overexpression of Tankyrase 1 and 2 in telomerase positive cell lines resulted in telomere elongation by releasing Trf1 from telomeres and by

regulating the access of telomerase to telomeres (Cook *et al.*, 2002). Telomere length regulation by Tankyrase 1 and 2 is probably not conserved in mice, where knockouts of either tankyrase 1 or 2 do not have any effect on telomere length or capping (Chiang *et al.*, 2006; Hsiao *et al.*, 2006; Muramatsu *et al.*, 2007).

1.3.2 T-loop formation is regulated by shelterin

Protection of telomeres is important to maintain the stability of chromosomes. In this context, the shelterin complex is thought to protect the telomeres by facilitating the formation of telomere secondary structures known as t-loops. The formation of t-loops originates from the invasion of the 3' single-stranded overhang to double-stranded telomeric DNA (Figure 1.3). These structures were first identified by electron microscopy in mouse and human cells extracts, where a wide variation of t-loop sizes, from very small (~1 kb) to very large (~25 kb) were observed. *In vitro* formation of t-loops was induced when a linear telomeric DNA molecule was incubated with the Trf2 protein; however incubation with Trf1 or the absence of telomeric proteins did not form t-loop structures and this suggested that Trf2 was directly involved in the formation of t-loops (Griffith *et al.*, 1999).

The presence of the 3' overhang is essential for t-loop formation. *In vitro* experiments have shown that blunt ends or 5' overhangs do not generate t-loops. On the other hand, the length of the 3' overhang is not critical for the formation of these structures since one telomere repeat was sufficient for t-loop formation (Stansel *et al.*, 2001). It is believed that proteins involved in recombination/repair that bind Trf2, such as Mre11 and Blm helicase can participate in the formation of t-loops *in vivo* by promoting the unwinding of DNA to facilitate the invasion of the single-stranded into double-stranded DNA. Furthermore, Trf2 protein was found to induce changes in the conformation of DNA by generating positive supercoiling. It is therefore proposed that this event can induce relaxation and untwisting of the DNA immediately adjacent to the Trf2-supercoil DNA complex thus promoting strand invasion and t-loop formation (Amiard *et al.*, 2007). Using electron microscopy techniques, it was observed that Trf2 localised preferentially to the loop junction, reinforcing the idea that Trf2 is important in t-loop formation (Stansel *et al.*, 2001).

Apart from physically protecting the telomeres against nuclease degradation, t-loops may be involved in the control of telomere length by regulating the access of telomerase to the 3' overhang end, thus controlling telomere length homeostasis (discussed in the next section).

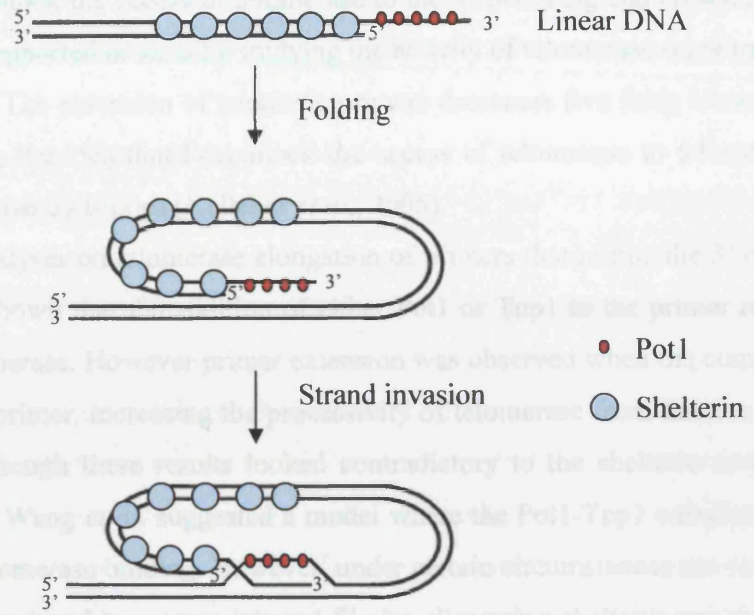


Figure 1.3 Model for the formation of t-loops mediated by shelterin. The shelterin proteins Trf1 and Trf2 promote the folding of linear DNA. Trf2 protein is involved in the generation of t-loops and it is found at the junction between double and single-strand DNA. Although Pot1 forms part of the shelterin complex, it has been represented separately to indicate its possible role in protecting single-strand DNA once the t-loop is formed (Figure modified from de Lange, 2005).

1.3.3 Telomere length regulation by shelterin

All members of the shelterin complex are involved in controlling the synthesis of telomeric DNA by regulating the access of the enzyme telomerase to telomeres. It has been proposed that shelterin is a *cis*-acting element that inhibits telomeric DNA elongation by telomerase in cells that have long telomeres. However, in cells with short telomeres, which contain less shelterin, telomerase can replenish telomere length (Smogorzewska and de Lange, 2004).

In general, it has been observed that depletion or mutation of individual components of shelterin affect telomere length. For examples expression of a mutant form of Pot1 (ΔOB), that does not bind single-stranded DNA and that represses the expression of endogenous Pot1, results in telomere elongation of about 230 bp per population doubling. Thus suggesting that

Pot1 can directly block the access of telomerase to the 3' overhang end (Loayza and de Lange, 2003). This was supported *in vitro* by studying the activity of telomerase on primers that contain telomeric repeats. The extension of telomeric primers decreases five folds when Pot1 protein is present supporting the idea that Pot1 block the access of telomerase to telomeres which may also be helped *in vivo* by t-loops (Kelleher *et al.*, 2005).

Further analyses on telomerase elongation of primers that mimic the 3' end overhang of a telomere have shown that the addition of either Pot1 or Tpp1 to the primer reaction blocked the action of telomerase. However primer extension was observed when the complex Pot1-Tpp1 was added to the primer, increasing the processivity of telomerase from three to six fold (Wang *et al.*, 2007). Although these results looked contradictory to the shelterin model of telomere length regulation, Wang *et al.* suggested a model where the Pot1-Tpp1 complex protects the 3' overhang from telomerase binding. However, under certain circumstances not yet identified, this complex can be removed to a more internal 5' site, disrupting shelterin and then allowing the extension of telomeres.

The effect of other members of shelterin on telomere length regulation has also been studied. In humans, overexpression of Trf1 in a telomerase positive cell line (HT1080) induced gradual shortening of telomeres. On the contrary, telomere elongation was achieved by the over expression of a dominant-negative Trf1 that inhibited the binding of the endogenous Trf1 to telomeres (van Steensel and de Lange, 1997). Similar to Trf1, Trf2 overexpression induced progressive shortening of telomeres in telomerase positive cell lines without affecting the expression of telomerase. This implied that the regulation of telomere length by the Trf1 and Trf2 proteins was mediated by access of telomerase to the telomeres (probably via a secondary telomere structures such as t-loops) but not by affecting telomerase expression itself (Smogorzewska *et al.*, 2000). Apart from controlling telomere length, Trf2 (but not Trf1) is involved in preventing end-to-end fusions between chromosomes. A dominant negative allele of *TRF2* was created to study the role of the Trf2 protein in telomere capping. The dominant negative allele contains a deletion in the C-terminal Myb domain but retains the homodimer formation domain (*TRF2*$^{\Delta B\Delta M}$). When the *TRF2*$^{\Delta B\Delta M}$ mutant allele is expressed in cells, it can bind to endogenous Trf2 proteins so preventing the binding of the resulting heterodimer to telomeres. The expression of this mutant allele led to the generation of end-to-end fusions observed as an increase in dicentric and multicentric chromosomes, most likely induced by non-homologous end joining (NHEJ) (van Steensel *et al.*, 1998). Furthermore, the 3' overhangs of

some telomeres were lost, suggesting that it was the DNA damage signal that induced NHEJ (further details are discussed below).

1.3.4 Epigenetic regulation of telomere length

The regulation of telomere length is not only mediated by shelterin but also by other factors that bind DNA or interact with the shelterin complex. In addition, chromatin modifications have important roles in telomere function and length regulation (reviewed in Blasco, 2007). Telomeric DNA contains nucleosomes with a similar composition to that found in heterochromatin. In fact, histone modification at telomeric chromatin produces changes in its organisation and its compaction, regulating telomere length (Gonzalo *et al.*, 2006). The main histone modification found at telomeric and subtelomeric chromatin is tri-methylation of histone 3 at lysine 9 and histone 4 at lysine 20 (H3K9 and H4K20). Histone acetylation which is normally associated with open chromatin conformation is not found at telomeric or subtelomeric chromatin. The importance of chromatin modification for telomere length regulation was demonstrated using mouse embryonic stem cells lacking histone methyltransferases (Suv39h1 and Suv39h2). The cells of these mice have reduced levels of methylation at H3K9 and H4K20 leading to defects in telomere length regulation and increasing telomere recombination events as seen by elongation of telomeres and an increase in telomere sister chromatid exchange events respectively (Garcia-Cao *et al.*, 2004; Benetti *et al.*, 2007). However, methylation of histones is not the only chromatin modification involved in telomere length regulation. It was found that methylation of CpG sequences has also an effect on telomere length. Although mice and human telomeres do not contain CpG sequences, they are found at subtelomeric regions. Mouse cells lacking methyltransferases such as Dnmt1, Dnmt3 and Dnmt3a have unmethylated CpG sites at subtelomeres. As a result of this loss, the telomeres become elongated, suggesting a role for subtelomeric chromatin organisation in controlling telomere length independent of the histone methylation status (Gonzalo *et al.*, 2006).

The control of telomere length by chromatin architecture has also been observed in fibroblast embryonic mouse cells that lacked telomerase (*Tert*$^{-/-}$). These cells showed telomere shortening due to lack of telomerase. However, once the telomeres were short, the heterochromatin markers such as H3K9, H4K20 and methylation of subtelomeric CpG were lost. Thus, it has been suggested that loss of telomeric repeats may lead to a more open state of the chromatin by the loss of heterochromatin markers (Benetti *et al.*, 2007).

1.3.4.1 Telomere position effect

Subtelomeric chromatin has other heterochromatic properties, it is replicated late in S phase, it has low gene density and it has the ability to silence nearby genes. This process of gene silencing is known as telomere position effect (TPE) and it was first observed in *Drosophila*, however most of the knowledge in this area comes from work carried out in *S. cerevisiae* (Gottschling *et al.*, 1990). In budding yeast, it was shown that genes inserted near the start of the telomere are repressed and this repression decreases as the genes are inserted further from the telomere (*i.e.* more centromeric). However TPE is discontinuous with the maximum repression found at the subtelomeric core X element but not in other region such as the Y' elements. In yeast, the silencing process is dependent on the telomeric binding protein Rap1 and the silent information regulators Sir1, Sir2 and Sir4 proteins. Sir proteins are recruited to telomeres by interactions with Rap1, forming repressive chromatin structures that restrict transcription of genes placed at these sites (Pryde and Louis, 1999). Furthermore, Sir proteins promote hypoacetylation of H3 and H4 causing formation of heterochromatin and transcriptional repression (Luo *et al.*, 2002).

In humans, TPE has also been observed and it has similar characteristics to that seen in yeast. For example, genes placed in telomeric regions show repression dependent on hypoacetylation (Baur *et al.*, 2001). Furthermore the repression is influenced by telomere length, *e.g.* elongation of telomeres by expression of telomerase enhances gene repression at telomeres whereas overexpression of Trf1 (that causes telomere shortening) reduces that effect. Thus the shelterin complex may also influence chromatin compaction at telomeres taking part in the telomere position effect process (Baur *et al.*, 2001; Koering *et al.*, 2002; Blasco, 2007).

1.4 Telomere maintenance mechanisms

In human somatic cells, telomeric sequences are lost by an average of 50 – 150 base pairs per cell cycle because of incomplete replication and nucleolytic processing at the chromosome termini (Harley *et al.*, 1990). It is thought that this shortening of telomeres limits the lifespan of the cells and acts as a tumour suppressor mechanism (Figure 1.4). However, cells have evolved mechanisms to overcome this limitation by activating telomere maintenance mechanisms resulting in the elongation of telomeric DNA. In most cancers, telomeric sequences are maintained by the activation of the enzyme telomerase. However, some cancers and approximately 35% of immortalised cells that are telomerase-negative use an alternative

lengthening of telomeres (ALT) mechanism (Figure 1.4) which is most probably based on recombination (Bryan *et al.*, 1997).

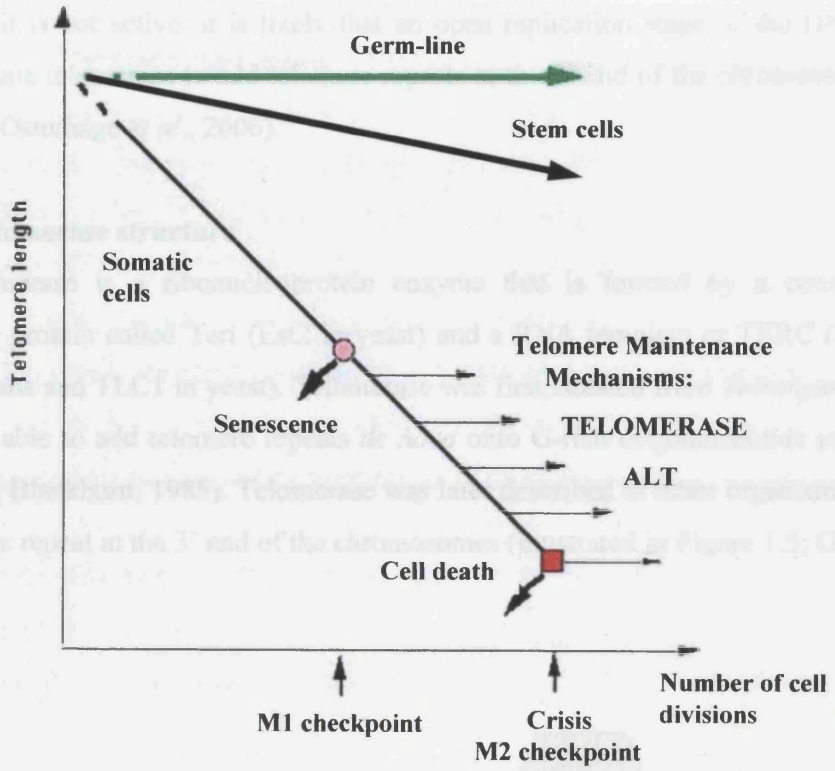


Figure 1.4 Replicative senescence and activation of telomere maintenance mechanisms. Somatic cells undergo replicative senescence generated by telomere shortening (M1 checkpoint). Some cells bypass the M1 checkpoint (*e.g.* by inactivation of *TB53* or *RB* gene) extending their life span until their telomeres reach a critical length where the cells die (M2 checkpoint). Rare cells can survive crisis by activating telomerase or ALT in order to stabilise their telomeres. Germ-line and stem cells maintain their telomeres via the expression of telomerase.

1.4.1 Telomere maintenance by telomerase

Expression of telomerase is the most common mechanism that tumour cells used to maintain the length of their telomeres. The expression of telomerase is cell cycle dependent. In *S. cerevisiae* it was shown that telomerase is active in late S phase and the beginning of G2, corresponding to the time telomeres are replicated. Since telomerase is recruited to telomeres in late G1 but it is not active, it is likely that an open replication stage of the DNA in S phase would facilitate telomerase to add telomere repeats at the 3' end of the chromosomes (Marcand *et al.*, 2000; Osterhage *et al.*, 2006).

1.4.1.1 Telomerase structure

Telomerase is a ribonucleoprotein enzyme that is formed by a conserved reverse transcriptase protein called Tert (Est2 in yeast) and a RNA template or TERC (also known as hTR in humans and TLC1 in yeast). Telomerase was first isolated from *Tetrahymena* and it was found to be able to add telomere repeats *de novo* onto G-rich oligonucleotide primers *in vitro* (Greider and Blackburn, 1985). Telomerase was later described in other organisms and found to add telomeric repeat at the 3' end of the chromosomes (illustrated in Figure 1.5; Greider, 1990).

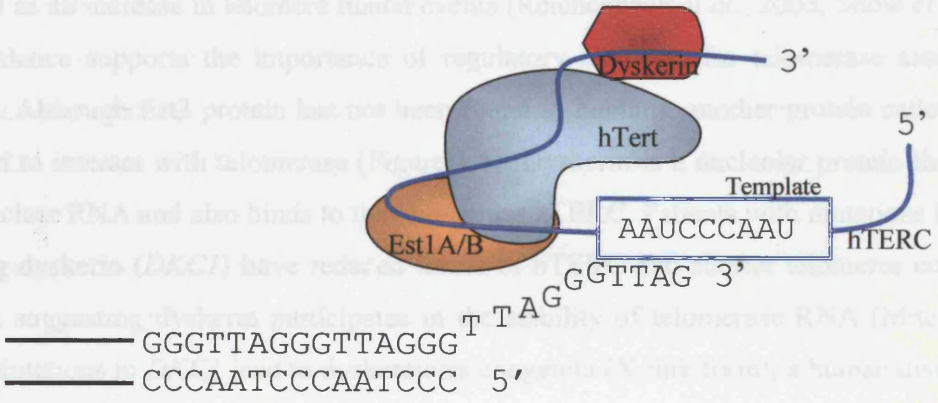


Figure 1.5 Elongation of telomeres by the enzyme telomerase. Telomerase uses the 3' overhang of the telomere as a primer for telomere extension and the template region of the TERC RNA is the substrate for reverse transcription. RNA binding proteins such as dyskerin and Est1 A/B are associated with telomerase and stabilise the RNA (hTERC) component of telomerase.

The catalytic component of telomerase, Tert, is highly conserved among organisms; however the RNA component, TERC, is variable in size and sequence. For example, the RNA particle varies from ~150 nt in ciliates to 400 to 600 nt in vertebrates and ~1300 nt in yeast (Chen and Greider, 2004). Despite this variation, the secondary structure of TERC contains conserved domains present from ciliates to vertebrates. These domains presumably serve in the interaction with the Tert component and thus are important for telomerase function (Chen *et al.*, 2000).

Some proteins have been found to associate with telomerase contributing to its function. For example, in *S. cerevisiae* Est1 and Est3 proteins are involved in the regulation of telomerase. Est1p has been found to interact with Cdc13 and together recruit telomerase to the telomeres (Lendvay *et al.*, 1996; Qi and Zakian, 2000). The function of Est1p is stressed when this protein is deleted. The result of this deletion is the shortening of telomeres and ultimately cell death. Recently, it was shown that the Est1 protein associates with telomeres in G1 phase, and this interaction was important for the recruitment of Est3 and telomerase assembly. However, the recruitment of telomerase by Est1 was insufficient to activate telomerase (Osterhage *et al.*, 2006).

In human cells, two homologs of Est1, hEst1A and hEst1B, had been found to associate directly with telomerase and contribute to its function (Figure 1.5). Additionally, overexpression of hEst1A protein in cancer-derived cells (HT1080 and HeLa) affected telomere capping, observed as an increase in telomere fusion events (Reichenbach *et al.*, 2003; Snow *et al.*, 2003). This evidence supports the importance of regulatory elements for telomerase assembly and function. Although Est3 protein has not been found in humans, another protein called dyskerin is known to interact with telomerase (Figure 1.5). Dyskerin is a nucleolar protein that binds to small nuclear RNA and also binds to the telomerase hTERC. Patients with mutations in the gene encoding dyskerin (*DKC1*) have reduced levels of hTERC and shorter telomeres compared to controls, suggesting dyskerin participates in the stability of telomerase RNA (Mitchell *et al.*, 1999). Mutations in *DKC1* lead to dyskeratosis congenita (X-link form), a human disease that is characterised by abnormal skin pigmentation, bone marrow failure and increased incidence of cancer (Dokal, 2000). Other mutations that cause dyskeratosis congenita include mutations in the *hTERC* gene (autosomal dominant form; Vulliamy *et al.*, 2001).

1.4.1.2 Regulation of telomerase expression

Telomerase is constitutively expressed in most unicellular organisms but it is tightly regulated in humans. The *TERC* gene encoding the hTERC RNA is expressed in the germ line, stem cells and differentiated somatic cells in humans. However, *hTERT* expression is tightly regulated and expression of this gene correlates with telomerase activity (reviewed in Smogorzewska and de Lange, 2004).

Telomerase is also strongly regulated during embryogenesis. During the first three months of human gestation, telomerase is expressed in most tissues. However, after that stage, telomerase activity is down-regulated or even switched-off and this coincides with tissue differentiation (Ulaner *et al.*, 1998). Telomerase activity is down-regulated after 11 weeks of gestation in heart and kidney, after 16 weeks in brain and 21 weeks in liver. In some organs such as the kidney, even though telomerase activity is detected only until week 12, hTERT transcripts are detected until week 21. However, these transcripts are variant-splice forms of the active hTERT mRNA. Thus, it is thought that telomerase activity can be regulated by either transcriptional control and/or alternative splicing of hTERT mRNA (Ulaner *et al.*, 1998). The regulation of *hTERT* during development might be mediated by transcription factors such as Myc and Sp1, since binding sites of these factors are found in the promoter of the *hTERT* gene. In the mouse, *in vivo* studies have shown that Myc regulates telomerase transcription in the skin. Myc is able to upregulate the expression of mouse *TERT* and as a consequence telomerase activity increases (Flores *et al.*, 2006). Growth factors and binding sites for steroid (such as oestrogen) have also been identified in the promoter of *hTERT* suggesting that some hormones also regulate telomerase activity (Soda *et al.*, 2000; Lebeau *et al.*, 2002). Similarly, the promoter of the *hTERC* gene contains binding sites for regulatory elements such as Sp1 and NF-Y. Furthermore, it has been found that the products of the retinoblastoma gene and the *MDM2* oncogene regulate *hTERC* expression and so act as positive and negative regulators of telomerase activity respectively (Zhao *et al.*, 2005).

1.4.1.3 Additional activities of telomerase

Telomerase was first described as a reverse transcriptase that stabilises the chromosome ends by adding new telomere repeats (Greider and Blackburn, 1985). However, this enzyme has been implicated in promoting cell division in epithelial cells by inducing the expression of

growth-promoting genes. Human mammary epithelial cells, ectopically expressing *hTERT*, exhibit upregulation in transcription of growth factors and downregulation in genes that suppress cell growth (Smith *et al.*, 2003).

Furthermore, it was shown by Stewart *et al.* (2002) that transfection of *H-RAS* oncogene in an immortalised cell line (GM847- SV40 transfected cell) that uses the ALT pathway (described in next section) do not readily form tumours when injected in nude mice. However, when the same cells were co-transfected with *hTERT*, they induced tumour formation even though the cells retained the ALT phenotype. Moreover, when the cells expressing the *H-RAS* oncogene were transfected with a mutant form of *hTERT* ($htert_{HA}$) that retains the catalytic activity but is unable to elongate telomeres, it induced tumour formation in the immunodeficient nude mice. Together, these studies suggest a role for telomerase in tumourigenesis by promoting cell growth that is independent of telomere length regulation (Stewart *et al.*, 2002).

Mouse cells have high levels of telomerase expression and contain long telomeres that persist over the life time of the animal. Thus, mouse cells enter senescence in a mechanism independent of telomere attrition. Nevertheless, overexpression of the catalytic component of telomerase, *mTERT*, in mouse basal keratinocytes, induced epidermal tumours by promoting cellular proliferation regardless of their long telomeres. Furthermore, these mice had an increase in wound healing capacity compared to wild-type mice, indicating again that an increase in cell proliferation occurs under *mTERT* over-expression (Gonzalez-Suarez *et al.*, 2001).

1.4.2 Alternative lengthening of telomeres

Although the activation of telomerase is the default telomere maintenance mechanism in some normal tissues and it is commonly activated in human tumours, some cells and tumours are able to maintain their telomeres by a telomerase independent mechanism. This process is called alternative lengthening of telomeres (ALT) and was first described in budding yeast. In humans it is believed that about 10% of tumours and 35% of *in vitro* immortalised cells use the ALT pathway to maintain telomere length (Bryan *et al.*, 1997; Henson *et al.*, 2002).

1.4.2.1 Alternative lengthening of telomeres in yeast

A yeast strain lacking the *EST1* gene (*est1*-Δ) shows progressive telomere shortening until the cells reach a critical point where they die. However, a small proportion of the yeast cells escape crisis and regain the ability to grow. These survivors display an increase number of

rearrangements at telomere and subtelomere regions generated by recombination events (Lundblad and Blackburn, 1993). The same *est1*-Δ phenotype is observed in strains that lack other genes that encode telomerase components such as *EST2*, *EST3* or *TLC1* (Singer and Gottschling, 1994; Lendvay *et al.*, 1996).

Telomerase negative survivors can be classified into two types according to their telomere composition; however both types are dependent on genes involved in recombination such as *RAD52*. Type I survivors maintain their telomeres by extensive amplifications of Y' elements, however the G-rich telomere repeats remain short. By contrast, type II survivors show reduced subtelomeric Y' amplifications but they contain very long and heterogeneous G-rich telomeric repeats of up to 12 kb. Another difference between both survivors is their growth rate. Type II survivors grow at the same rate as wild-type strains, however type I cells divide more slowly. It is thought that type I survivors duplicate their Y' elements through nonreciprocal translocation events promoted by the interstitial G-rich repeats that separate the Y' elements or by integration of extrachromosomal Y' elements into short telomeres (reviewed in Lundblad, 2002). In a similar way, type II survivors maintain their telomeres through recombination-based events. It has been proposed that elongation of telomeres is originated by unequal recombination between telomeres or by a break-induce replication event where the 3' end of a critical short telomere invade a second chromosome that acts as a template for telomere elongation (Teng and Zakian, 1999; McEachern and Haber, 2006).

Yeast strains with mutations in *RAD52* and telomerase die quickly and no survivors are recovered, demonstrating the total dependence on the homologous recombination *RAD52* gene (Lundblad and Blackburn, 1993). Other recombination genes such as *RAD50*, *RAD51*, *RAD54* and *RAD57* have been found to affect cell viability in the absence of telomerase but survivors do arise (Le *et al.*, 1997). However, a *rad50*, *rad51* and *tlc1* triple mutant strain does not generate survivors, suggesting *RAD50* and *RAD51* act in different pathways. In support of this idea, the function of Rad51 and Rad50 is linked to type I and type II survivors respectively (Chen *et al.*, 2001).

1.4.2.2 Alternative lengthening of telomeres in humans

Human cell lines displaying the ALT pathway are characterised by containing heterogeneous and long telomeres ranging from 2 up to 50 kb with an average length of ~20 kb (Murnane *et al.*, 1994; Bryan *et al.*, 1997). Using FISH analysis it has been observed that some

chromosomes have very long telomeres while others lack any detectable signal. Furthermore, ALT cells exhibit sudden gains or losses of telomere repeats sometimes involving many kb of sequences (Murnane *et al.*, 1994; Perrem *et al.*, 2001). This telomere length phenotype is not observed in telomerase positive cells, where cells have a more homogenous distribution of telomere length with an average length of around 10 kb (Counter *et al.*, 1992).

Another characteristic of ALT cells is the presence of doughnut-shaped ALT-associated PML bodies or APBs. These structures appear during immortalisation at the same time as ALT activation and they contain PML protein, telomeric DNA and telomere binding proteins such as Trf1 and Trf2. Apart from telomeric proteins, APBs also contain proteins involved in DNA repair, recombination and replication such as Rad52, Rad51, replication factor A (Yeager *et al.*, 1999), Wrn, Blm (Stavropoulos *et al.*, 2002) and the MRN complex (Mre11, Rad50, Nbs1; Zhu *et al.* 2000). The MRN complex is associated with APBs during late S phase of the cell cycle to G2 phase suggesting this complex together with recombination proteins may be involved in telomere maintenance by promoting recombination events (Wu *et al.*, 2000).

APBs have associated extrachromosomal telomeric DNA which can be found either in a circular or linear form (Tokutake *et al.*, 1998; Cesare and Griffith, 2004). Telomere extrachromosomal DNA is mainly found in the nucleoplasm and it is thought to participate in telomere maintenance in ALT cells. In *Kluyveromyces lactis*, circular DNA found in the cytoplasm of the cells is used for telomere elongation through a process known as gene conversion rolling-circle replication. This process was studied using artificial plasmids containing telomere repeats and the *URA3* marker gene. The circular plasmid was integrated into the telomeres of telomerase deficient cells but not in telomerase proficient *K. lactis*. Furthermore, plasmid integration was successful only when the plasmid was in a circular but not in a linear form. The use of extrachromosomal circle templates for telomere elongation was dependent on the *RAD52* gene. Thus it is proposed that a recombination process plays a role in the maintenance *via* extrachromosomal circles (Natarajan *et al.*, 2003).

1.4.2.3 Telomeric recombination in human ALT cells

ALT cell lines exhibit an increase in telomere recombination (*e.g.* telomere sister chromatid exchange; T-SCE) compared to telomerase positive cell lines. These recombination events occur between telomeric repeats and can be visualised using a technique called chromosome orientation FISH or CO-FISH (Londono-Vallejo *et al.*, 2004). CO-FISH depends

on the incorporation of bromodeoxyuridine (BrdU) during S phase, so that it is incorporated only into the newly synthesised DNA strand. Metaphase chromosomes are harvested and treated under UV light to nick the BrdU, which is then removed with exonuclease III. After that, the use of a strand specific telomere probe can indicate if a recombination event has occurred (according to the number of signals detected). Although multiple signals are usually interpreted as enriched for T-SCE, there is a limitation as this technique does not distinguish recombination events between sister chromatids, different chromosomes or between telomeres and extrachromosomal telomeric repeats.

Recombination has been proposed as one of the mechanisms used by ALT cells to elongate telomeres. However, analysis on the recombination levels in telomerase positive and ALT cells have shown similar levels at interstitial genomic sites, thus it has been suggested that an increase in recombination affects telomeres preferentially (Bechter *et al.*, 2003, 2004a). However, there is evidence of increased instability in a minisatellite (MS32) in ALT cell lines and tumours but not found in telomerase positive cells. MS32 is a GC-rich minisatellite composed of tandem arrays of 29-bp units located on chromosome 1q. Other minisatellites have also been screened for increased instability, though instability has only been found at the MS32 minisatellite. These findings suggest that ALT cells have elevated recombination, which might be limited to some loci (Jeyapalan *et al.*, 2005).

Further evidence suggests that intertelomeric recombination may underlie the elongation of telomeres in ALT cells. An ALT cell line (GM847) was transfected with a plasmid containing telomere repeats and after several population doublings, the plasmid was integrated and propagated (presumably by homologous recombination) to as many as 10 chromosome ends (Dunham *et al.*, 2000). This evidence has led to the proposal of several models of telomere maintenance in ALT (Figure 1.6). All of them involve the invasion of the 3' overhang of the telomere to a double-strand DNA that serves as a template for elongation. The double-strand DNA is thought to be either the same telomere through the formation of a t-loop structure, another telomere or extrachromosomal DNA such as that found in APBs (Henson *et al.*, 2002). Additionally, recombination proteins found in APBs are thought to be involved in telomere maintenance in ALT cells.

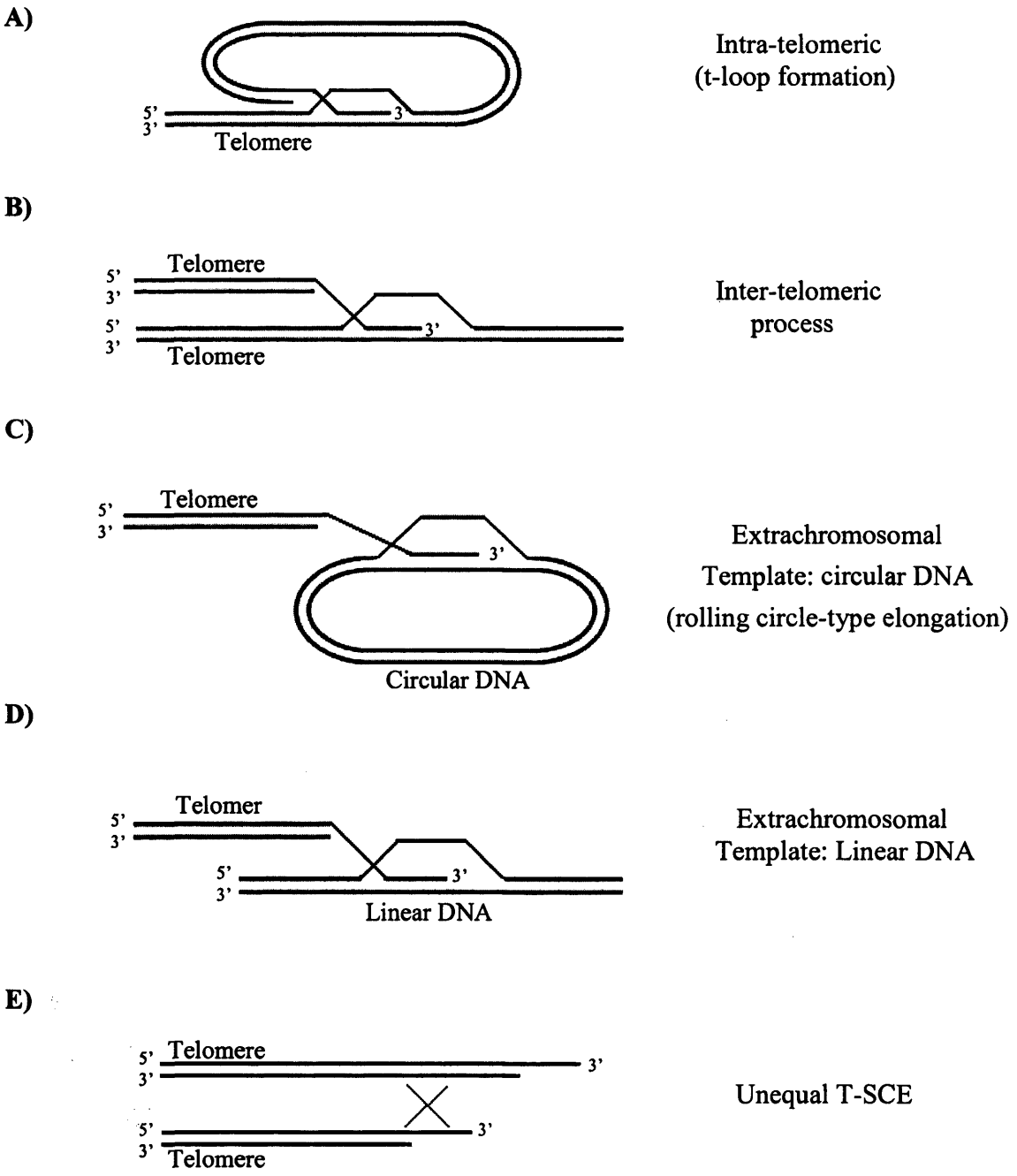


Figure 1.6 Models for telomere maintenance used by ALT cells. The 3' end of the telomere invades a double strand DNA which is used as a template for telomere elongation. The template could be the same telomere (A), another telomere (B) or extrachromosomal DNA either in linear (C) or circular form (D). Unequal telomere sister chromatid exchange (T-SCE) is elevated in ALT; however it is unlikely to give a net gain of telomere length (E). Figure modified from Henson *et al* (2002).

Direct evidence of a recombination-like mechanism acting at telomeres in ALT cells was demonstrated by mapping the interspersion pattern of telomere-variant repeats using a PCR method. By analysing the first 1 kb of individual telomeres, it was shown that some clones from different ALT cells contained novel telomeres that were a composite of the progenitor telomere fused to a different telomere repeat array. These mutant telomeres most likely arose by inter-molecular process and they were not found in telomerase positive or primary cell lines (Varley *et al.*, 2002).

1.4.2.4 Coexistence of ALT and telomerase activation

The ALT phenotype is characterised by the absence of detectable telomerase activity due to lack of *hTERC* and/or *hTERT* expression. The lack of telomerase activity can be produced by epigenetic modifications in the promoter of *hTERC* and *hTERT*; for instance, methylation of lysine 9 of histone H3 was found in ALT cells but not in telomerase positive cells (Atkinson *et al.*, 2005). However, expression of telomerase in ALT cell lines by expression of exogenous *hTERT* does not repress the ALT phenotype, though it induces some changes. For example, shorter telomeres normally found in ALT cells became elongated suggesting telomerase acts in the shorter telomeres. Nonetheless some other hallmarks of ALT cells such as very long telomeres, APBs and extrachromosomal circles persist even though telomerase was activated; indicating that telomerase itself is not enough to repress the ALT pathway (Grobelny *et al.*, 2001; Perrem *et al.*, 2001) and showing that these two mechanisms can coexist within a cell.

Telomerase activation and the ALT pathway are the two main mechanisms that SV40 immortalised cell lines and tumours use to maintain their telomeres. However, some cell lines have been found to lack both these mechanisms (Cerone *et al.*, 2005; Fasching *et al.*, 2005; Marciniak *et al.*, 2005). In tumours such as sarcomas there is a high incidence of ALT activation but also a high proportion of these tumours do not show any known mechanism for telomere maintenance. Thus, it is possible that more than one ALT mechanism exists in tumours or some tumours lack a telomere maintenance mechanism (Ulaner *et al.*, 2003; Henson *et al.*, 2005).

1.4.3 Telomere maintenance *via* retrotransposition events in *Drosophila melanogaster*

As discussed above, telomerase and ALT are the main pathways that eukaryotes use to maintain their telomere length; however, *D. melanogaster* uses retrotransposition events to maintain its telomeres (Figure 1.7). *Drosophila* telomeres lack the 3' overhangs and they are not formed by G-rich sequences, instead its telomeres are composed of three classes of related LINE-like retrotransposons (Het-A, TART and TAHRE). These telomeric retrotransposons are composed of a long 3' UTR region, a reverse transcriptase gene (only found in TART and TAHRE) and a *GAG* gene (Abad *et al.*, 2004; Pardue *et al.*, 2005). The products of *GAG*, Gag proteins, are involved in transporting the retrotransposon RNA to the nucleus and in delivering it to telomeres (Rashkova *et al.*, 2002).

Telomeric retrotransposons are reverse transcribed at the integration sites binding to telomeres through their 3' poly(A) tail and from that binding, they prime reverse transcription of the RNA directly onto the chromosome ends to generate the first DNA strand using the free 3' hydroxyl group (Figure 1.7). The second strand of DNA is synthesised by DNA polymerases which may be dependent on DNA repair enzymes, however the details of this process are not known (Walter *et al.*, 2007). The telomeric retrotransposons vary in number and in type depending on the telomere analysed, however, they are organised as head-to-tail arrays. Their presence at telomeres does not follow any pattern, suggesting retrotransposition event occurs randomly (Pardue *et al.*, 2005).

Although there are differences in the composition and maintenance of *Drosophila* telomeres, their function is conserved. They protect chromosomes from non-homologous end joining helped by proteins such as the heterochromatin protein 1 (Hp1), Atm and the Mre11/Rad50 complex (Bi *et al.*, 2004). Studies in *Drosophila atm* and *mre11* knockouts showed a significant increase in telomere fusions, chromosome breakage and rearrangements. Furthermore, *mre11* and *atm* mutants had a reduction in the number of HeT-A elements at telomeres, suggesting telomere maintenance is dependent of *ATM* and *MRE11* genes (Bi *et al.*, 2004).

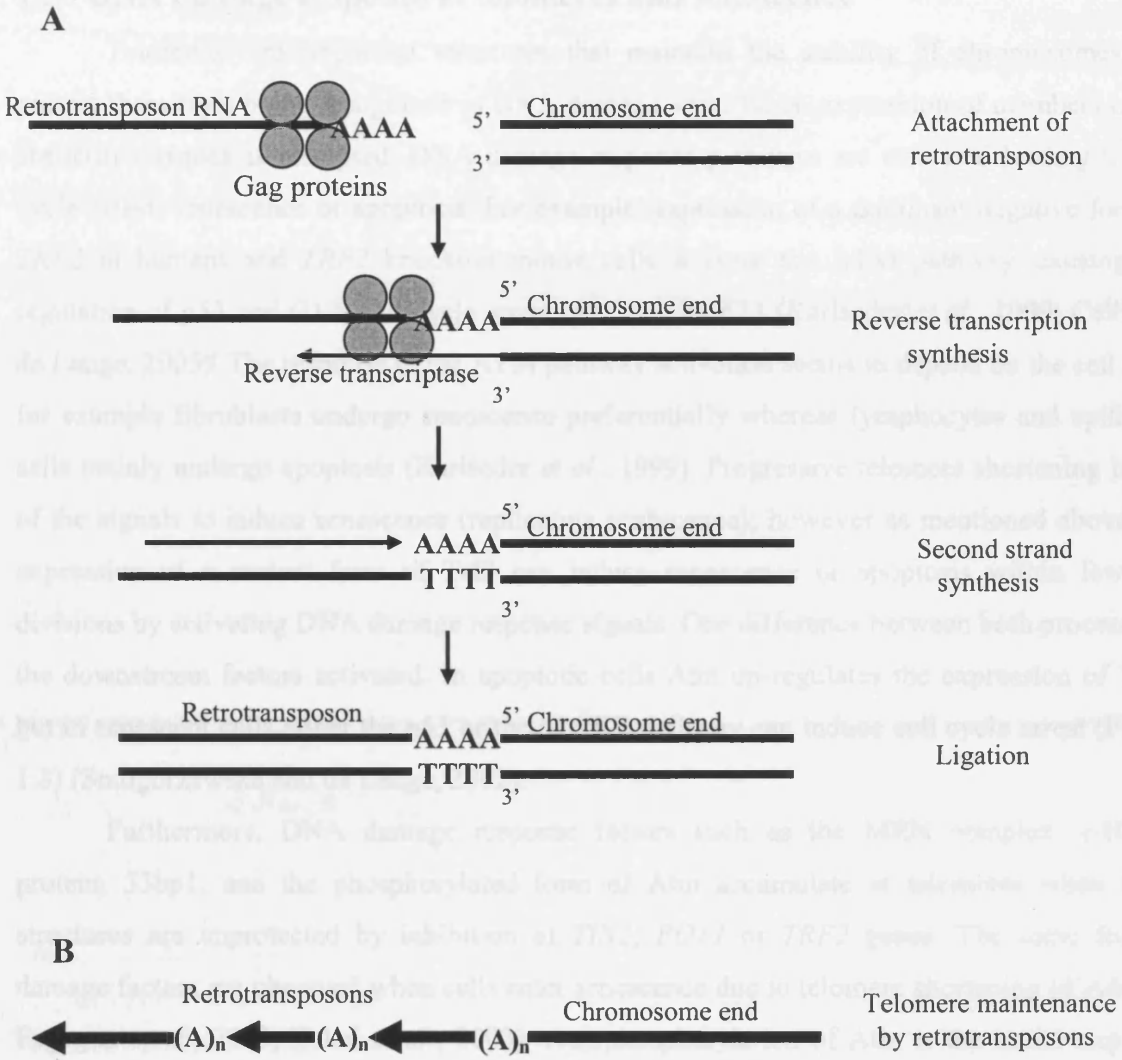


Figure 1.7 Telomere maintenance by retrotransposition events in *Drosophila*. Retrotransposon RNA binds to the 3' end of the telomeres through its poly(A) tail. The Gag proteins are required for the targeting of retrotransposons to the end of the chromosome (*A*). Reverse transcriptase synthesises the complementary RNA strand, while a second strand synthesis takes place to replace the RNA sequence. More than one retrotransposon can be found at a single telomere, organised as head-to-tail arrays with their poly(A) tail facing towards the centromere (*B*). Figure modified from Mason and Biessmann, 1995.

1.5 DNA damage response at telomeres and senescence

Telomeres are important structures that maintain the stability of chromosomes and protect them from being recognised as DNA damage sites. When expression of members of the shelterin complex is disrupted, DNA damage response pathways are activated leading to cell cycle arrest, senescence or apoptosis. For example, expression of a dominant negative form of *TRF2* in humans and *TRF2* knockout mouse cells activate the ATM pathway, causing up-regulation of p53 and G1/S cell cycle arrest mediated by p21 (Karlseder *et al.*, 1999; Celli and de Lange, 2005). The response to the ATM pathway activation seems to depend on the cell type, for example fibroblasts undergo senescence preferentially whereas lymphocytes and epithelial cells mainly undergo apoptosis (Karlseder *et al.*, 1999). Progressive telomere shortening is one of the signals to induce senescence (replicative senescence); however as mentioned above, the expression of a mutant form of Trf2 can induce senescence or apoptosis within few cell divisions by activating DNA damage response signals. One difference between both processes is the downstream factors activated. In apoptotic cells Atm up-regulates the expression of *TP53* but in senescent cells either the p53 or the p16/RB pathway can induce cell cycle arrest (Figure 1.8) (Smogorzewska and de Lange, 2002).

Furthermore, DNA damage response factors such as the MRN complex, γ-H2AX protein, 53bp1, and the phosphorylated form of Atm accumulate at telomeres when these structures are unprotected by inhibition of *TIN2*, *POT1* or *TRF2* genes. The same foci of damage factors are observed when cells enter senescence due to telomere shortening (d'Adda di Fagagna *et al.*, 2003; Takai *et al.*, 2003). Autophosphorylation of Atm is the initial step that activates the ATM pathway. Interestingly, *in vitro* studies identified interactions of Trf2 with the Atm serine residue that is autophosphorylated after DNA damage response. Thus, it is thought that Trf2 protects the telomeres from being recognised as DNA damage regions by inhibiting Atm autophosphorylation (Karlseder *et al.*, 2004).

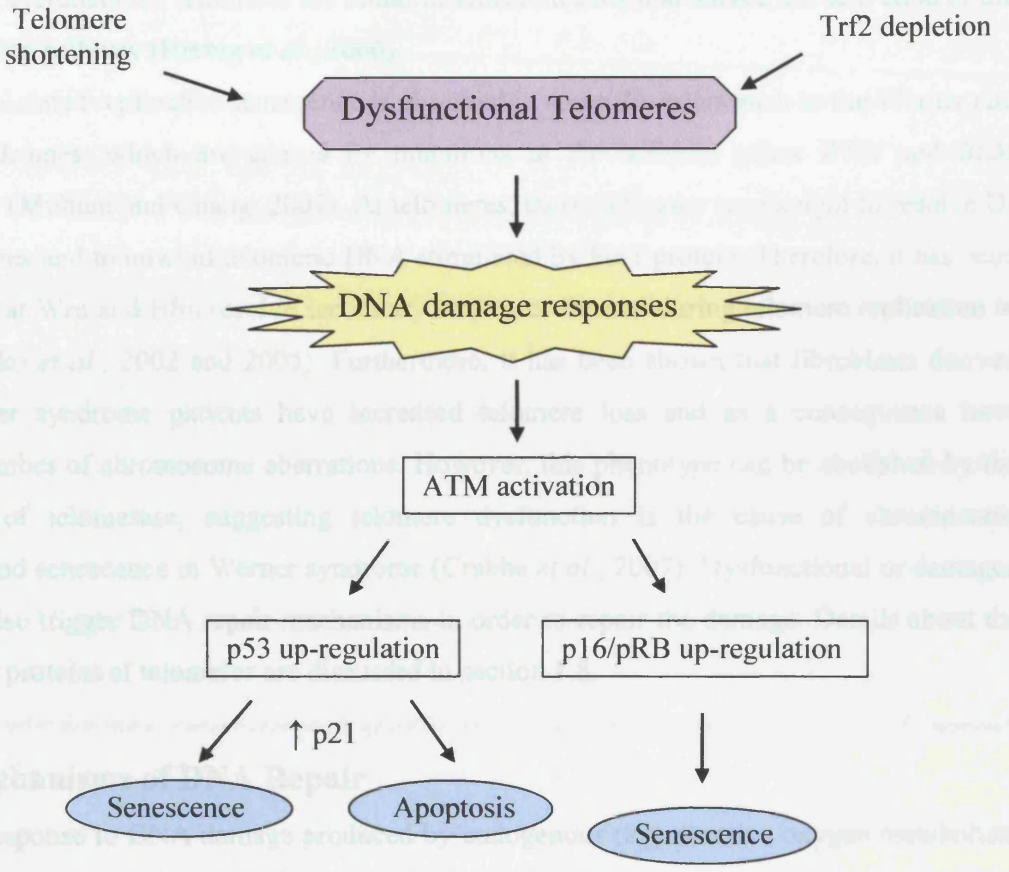


Figure 1.8 Senescence and apoptosis induced by dysfunctional telomeres. Telomere shortening or disruption of the shelterin complex (for example by depletion of Trf2 protein) can lead to dysfunctional telomeres that activate DNA damage signals. Atm autophosphorylation activates either the p53 or RB pathway resulting in senescence or apoptosis.

Dysfunctional telomeres have been identified *in vivo* in senescent cells of human tissue of aged liver, atherosclerotic plaques and skin. The identification of senescent cells in human tissue has been carried out by the presence of the senescence-associated β-galactosidase staining, one of the markers used to distinguish senescence cells (Dimri *et al.*, 1995; Campisi, 2005). Furthermore, senescent cells have been identified in aging primates. In fact, aged baboon skin fibroblasts contained DNA damage foci formed by the γ-H2AX, p53 binding protein (53bp1) and Atm kinase. These foci co-localised with telomeres and were increased exponentially with age, reaching 15% to 20% of all skin cells in very old baboons. These results

suggest that dysfunctional telomeres are found in senescent cells and induce the activation of the Atm signalling pathway (Herbig *et al.*, 2006).

Accelerated replicative senescence is observed in some disorders such as the Werner and Bloom syndromes, which are caused by mutations in the helicase genes *WRN* and *BLM* respectively (Multani and Chang, 2007). At telomeres, these helicases are thought to resolve D-loop structures and to unwind telomeric DNA stimulated by Pot1 protein. Therefore, it has been suggested that Wrn and Blm resolve secondary structures formed during telomere replication *in vivo* (Opresko *et al.*, 2002 and 2005). Furthermore, it has been shown that fibroblasts derived from Werner syndrome patients have increased telomere loss and as a consequence have elevated number of chromosome aberrations. However, this phenotype can be abolished by the expression of telomerase, suggesting telomere dysfunction is the cause of chromosome instability and senescence in Werner syndrome (Crabbe *et al.*, 2007). Dysfunctional or damaged telomeres also trigger DNA repair mechanisms in order to repair the damage. Details about the DNA repair proteins at telomeres are discussed in section 1.8.

1.6 Mechanisms of DNA Repair

In response to DNA damage produced by endogenous (*e.g.* reactive oxygen metabolism products) or exogenous (*e.g.* UV, ionizing radiation and reactive chemicals) sources, cells have evolved DNA repair mechanisms to correct the damage and to maintain chromosome integrity. However, if there is excess of DNA damage, cells stop dividing (cell cycle arrest or senescence) or they die through apoptosis. DNA repair pathways correct the damage and therefore maintain the integrity of the genome, including telomeres. In the next sections, the main pathways and the proteins involved in DNA repair are discussed. In addition, the involvement of these pathways with telomere stability is considered.

There are at least five major DNA repair pathways that maintain genome integrity in humans. The DNA double-strand breaks are either repaired by the homologous recombination repair (HRR) or the non-homologous end joining (NHEJ) pathways. The repair of damaged or mis-incorporated bases in DNA is achieved by nucleotide excision repair (NER), base excision repair (BER) or mismatch repair (MMR) pathways. These three mechanisms act in a similar way; they remove the damaged bases, insert new bases to fill the gap created and ligate the newly incorporated bases into the DNA strand. Although these mechanisms use similar steps to repair the damage, they recognise different DNA damage using specific proteins and enzymes.

1.6.1 DNA double-strand break repair

DNA double-strand breaks (DSBs) can arise by DNA damage agents (generated by exogenous or endogenous agents) or when DNA replication forks encounter a DNA lesion *e.g.* DNA single-strand breaks. Defects in the DSB repair pathway can generate chromosome aberrations such as translocations, deletions or inversions. These alterations can lead to loss of growth control or cell death (van Gent *et al.*, 2001).

1.6.1.1 Homologous recombination repair (HRR)

HRR is a high fidelity pathway to repair DSBs that uses homologous sequences as a template to re-synthesise damaged or missing information at the break site. However genomic instability can arise if HRR is not properly regulated. The HRR process requires a 3' single-stranded tail or overhang in the broken chromosome which is generated by 3' to 5' resection at the DSB by the Mre11/Rad50/Nbs1 complex (Paull and Gellert, 1998). The generated 3' overhang is coated with the replication protein A (RPA) and Rad52 which protects against exonucleolytic degradation (Stasiak *et al.*, 2000). Rad51 is the main player in HRR, facilitating the search for homologous sequences and strand invasion (Baumann *et al.*, 1996). Rad51 is helped by a series of mediator proteins such as the Rad51 paralogues (Xrcc2, Xrcc3, Rad51B, Rad51C, Rad51D), Rad52 and Brca2. Among these proteins, Brca2 has been shown to regulate the transport of Rad51 into the nucleus and so mediates Rad51 binding to the DSB (Davies *et al.*, 2001).

The process of strand invasion in search for homologous sequences displaces one strand of the template DNA forming a D-loop structure (Figure 1.9). Once the damaged strand is paired with a homologous sequence, DNA synthesis takes place. In most of the cases, the template strand is the sister chromatid and the repair process is mediated by a gene conversion event. After DNA synthesis has taken place, the invading strand is released by branch migration of the Holliday junction mediated by proteins such as Rad54, Wrn and Blm. These proteins have the ability to dissociate any secondary structure formed during replication therefore they can prevent the generation of aberrant recombination products. Finally, the remaining gaps are filled by polymerases and DNA ligase I closes residual nicks (Constantinou *et al.*, 2000; Karow *et al.*,

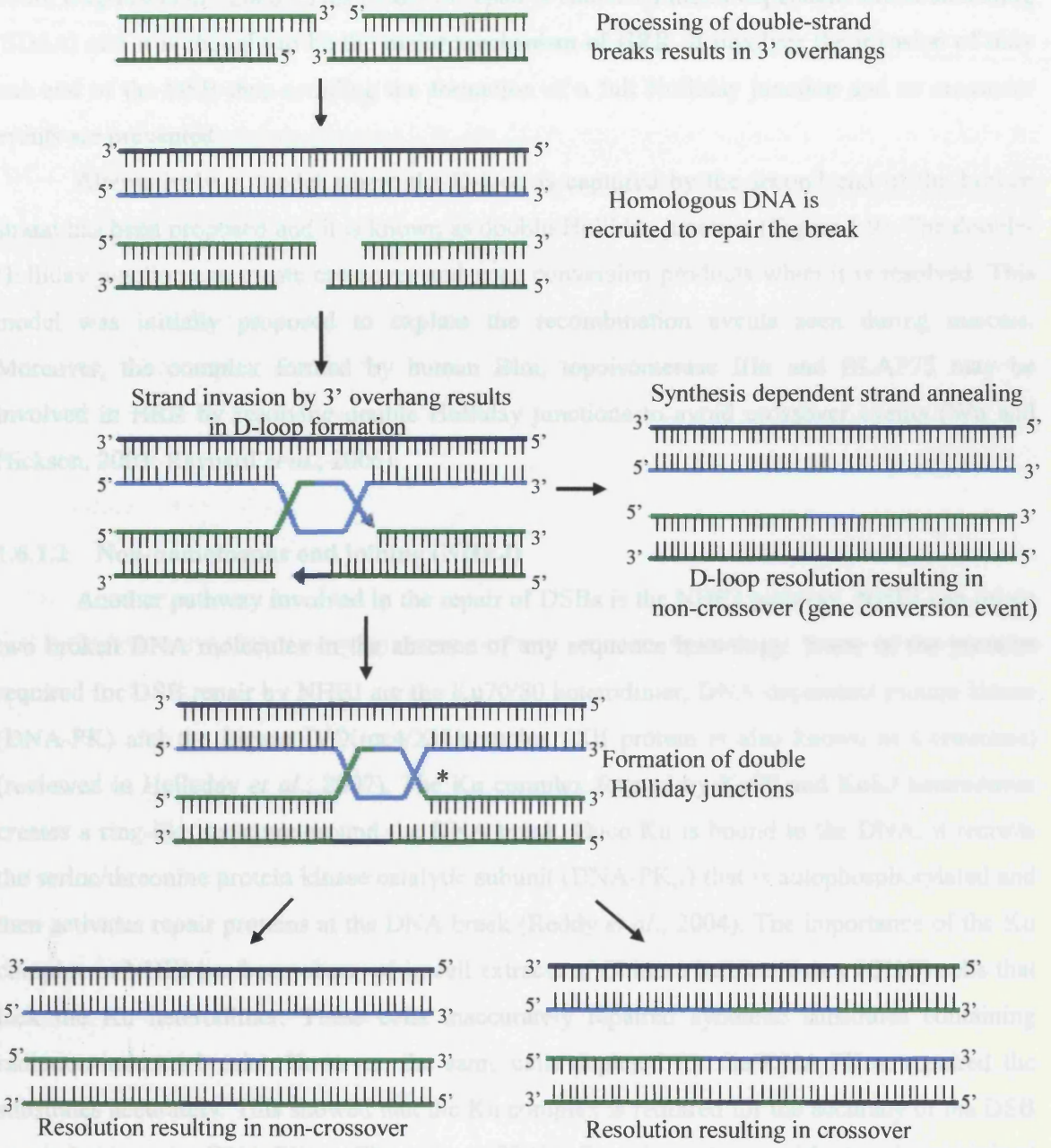


Figure 1.9 DSB repair by homologous recombination. The repair process initiates by the resection of the DSB by the Mre11/Rad50/Nbs1 complex generating a 3' overhang to which Rad52 and RPA bind. Rad51 protein provokes DNA strand displacement with the undamaged homologous DNA molecule (D-loop formation). D-loop structures are resolved by Wrn, Blm and Rad54 proteins. However, the D-loop can be captured by the second 5' end of the break (*) leading to the formation of a double Holliday junction structure. Resolution of this structure can generate crossover products.

2000; Bugreev *et al.*, 2006). This model of repair is called synthesis-dependent strand annealing (SDSA) and it is thought to be the major mechanism of HRR. It involves the invasion of only one end of the DSB thus avoiding the formation of a full Holliday junction and so crossover events are prevented.

Alternatively, a model where the D-loop is captured by the second end of the broken strand has been proposed and it is known as double-Holliday junction (Figure 1.9). The double-Holliday junction can create crossover and gene conversion products when it is resolved. This model was initially proposed to explain the recombination events seen during meiosis. Moreover, the complex formed by human Blm, topoisomerase IIIα and BLAP75 may be involved in HRR by resolving double Holliday junctions to avoid crossover events (Wu and Hickson, 2003; Raynard *et al.*, 2006).

1.6.1.2 Non-homologous end joining (NHEJ)

Another pathway involved in the repair of DSBs is the NHEJ pathway. NHEJ can rejoin two broken DNA molecules in the absence of any sequence homology. Some of the proteins required for DSB repair by NHEJ are the Ku70/80 heterodimer, DNA-dependent protein kinase (DNA-PK) and the Ligase IV/Xrcc4/Xlf complex (Xlf protein is also known as Cernunnos) (reviewed in Helleday *et al.*, 2007). The Ku complex formed by Ku70 and Ku80 heterodimer creates a ring-like structure around the DNA break. Once Ku is bound to the DNA, it recruits the serine/threonine protein kinase catalytic subunit (DNA-PK_{cs}) that is autophosphorylated and then activates repair proteins at the DNA break (Reddy *et al.*, 2004). The importance of the Ku complex in NHEJ has been observed in cell extracts of Chinese hamster ovary (CHO) cells that lack the Ku heterodimer. These cells inaccurately repaired synthetic substrates containing radiation-induced breaks. However, the same cells depleted for the DNA-PKcs, repaired the substrates accurately. This showed that the Ku complex is required for the accuracy of the DSB repair but not the DNA-PKcs (Chen *et al.*, 2001). Complementary or blunt ends are joined together by a complex formed by Ligase IV and the Xrcc4/Xlf complex (Ramsden and Gellert, 1998). The Xrcc4/Xlf complex may be required for targeting Ligase IV to the DNA ends since it can interact with the Ku complex and the DNA-PK_{cs} (Ahnesorg *et al.*, 2006, Buck *et al.*, 2006). The complex formed by Ligase IV/Xrcc4 and Ku is highly conserved among eukaryotes, however DNA-PK_{cs} is absent from some organisms such as *S. cerevisiae*.

DNA end-processing enzymes such as Wrn, Artemis, Tdp1 and polynucleotide kinases are required prior to repair to make the DNA ends compatible for ligation. For example, DSBs induced by ionizing radiation may contain gaps, 3'- phosphate or phosphoglycolate and 5'-hydroxyl groups. These modifications in the DNA may create molecules that are unable to ligate. Therefore, enzymes such as Wrn that contain a 3'-5' exonuclease activity can eliminate those groups. In addition, Wrn has been found to interact with DNA-PK_{cs} and is recruited to DNA by the Ku complex (Li and Comai, 2001).

It is not clear how a cell decides between the use of HRR or NHEJ after a double-strand break. However these two repair pathways are cell cycle dependent, thus the choice of pathway may depend on the cell cycle stage at which the DSB is identified. HR plays a major role during the S/G2 phases when sister chromatids are paired whereas NHEJ is mainly active in G1/early S phase (Hendrickson, 1997).

1.6.2 Base excision repair (BER)

Endogenous or exogenous processes that damage DNA can produce apurinic/apyrimidinic sites, methylations, deaminations, alkylations and oxidations of DNA bases. These base modifications can be recognised by the BER pathway. The process of repair mediated by BER is initiated by the recognition and removal of the damage by DNA glycosylases, which cleave the N-glycosyl bond between the modified base and the sugar of the DNA backbone resulting in apurinic/apyrimidinic (AP) sites (reviewed in Dizdaroglu, 2005). Several glycosylases have been found in humans and some of them catalyse the N-glycosyl cleavage of specific base modifications, for example, some glycosylases are specific for oxidatively modified bases (homologues of *E. coli* endonuclease III), alkylated bases (methylpurine-DNA glycosylases) or for the removal of uracil (uracil-DNA glycosylases).

Once the modified base has been removed by a DNA glycosylase/AP lyase, the gap is filled by DNA polymerase β and the nick is ligated by the Ligase III/Xrcc1 complex (Matsumoto and Kim, 1995). Sometimes, the removal of the damage involves more than one nucleotide and this is called long-patch BER. Long-patch BER shares the same characteristics described for single base repair or short-patch BER regarding the recognition and cleavage of the damage bases. However, long-patch BER is dependent on the proliferating cell nuclear antigen (PCNA) and the replication factor C (RFC) which regulates the polymerase and ligase activity. Strand displacement of the DNA region containing the damage is achieved by Polδ/ε,

PCNA and finally the flap generated (2 - 10 nt) is removed by Fen1. The patch created is then synthesised by Polδ/ε helped by the PCNA protein and ligated by Ligase I (Matsumoto *et al.* 1999; Sattler *et al.*, 2003).

1.6.3 Nucleotide excision repair (NER)

NER is a repair pathway that is mainly involved in the correction of UV-induced photoproducts (*e.g.* cyclobutane pyrimidine dimers -CPDs- and 6-4-photoproducts) and damage produced by chemotherapeutic agents that distort the DNA double helix structure. NER is divided in two sub-pathways, the global genomic repair (GGR) engaged in the repair of lesions present everywhere in the genome and the transcription-coupled repair (TCR) involved in the repair of lesions present on the transcribed strand of transcriptionally active genes. Both sub-pathways share the same sequence of events: recognition of the damage, incisions on either side of the damage, excision of the damaged nucleotides by removing a 24-32 nucleotides patch, DNA synthesis using the complementary strand and ligation to close the final nicks (Aboussekhra *et al.*, 1995; Araujo *et al.*, 2000). More than 30 proteins are involved in NER and are shared between the two sub-pathways. However, some of the proteins used for the recognition of the damage and the repair rate in each sub-pathway are different.

Lesions generated in the transcribed-strand of active genes are repaired by TCR more quickly than lesions found in the non-transcribed strand or in other genomic sequences, which are repaired by GGR. In the TCR pathway, the DNA damage located in the transcribed strand blocks the RNA polymerase II transcription and this activates the DNA repair responses. The first proteins to be recruited to the damage are the ATP dependent Csa and Csb (Rad28 and Rad26 in yeast, respectively) which release the RNA polymerase and recruit further repair proteins such as the XPG and XPF nucleotide excision repair proteins and the TFIIH repair complex containing helicases XPB and XPD (Laine and Egly, 2006). In contrast to TCR, GGR is initiated by the recognition of lesions by the damage-recognition proteins XPC-HR23B instead of being recognised by stalled-transcription complexes. *XPC* deficient cells are defective for GGR but not for TCR, supporting the idea that the XPC-HR23B is specific for GGR. Subsequently, XPA is recruited to the damage and the repair process continues in a similar manner to that seen in TCR (Volker *et al.*, 2001). Mutations in *XPB*, *XPD* or *XPG* genes are associated with Xeroderma pigmentosum, a syndrome that is characterised by high UV sensitivity and predisposition to skin cancer.

1. 7 Mismatch repair (MMR) pathway

Errors generated during the replication process such as misincorporation of nucleotides by DNA polymerases or the formation of extra-helical loops can be repaired by the MMR pathway. Furthermore, MMR is involved in the inhibition of recombination between non-identical sequences (homeologous recombination) and in promoting recombination events during meiosis. MMR proteins were initially described in prokaryotic organisms and their importance was demonstrated by the observation that defects in this pathway give rise to an elevation of spontaneous genomic mutations (reviewed in Kunkel and Erie, 2005).

1.7.1 Mismatch Repair in *Escherichia coli*

In *E. coli*, a series of proteins that includes helicases, exonucleases, DNA ligase and the MutS, MutL and MutH have been linked to the repair of DNA mismatches. The protein that initiates the MMR process is the MutS protein, which is involved in the recognition of base-base mismatches and small nucleotide insertion/deletion loops (Wang *et al.*, 2003). MutS has high binding affinity for mismatches, however its affinity depends on the composition of the mismatch, *e.g.* it has a higher affinity for G-T than for C-C mismatches (Lamers *et al.*, 2000). The MutS protein forms dimers that encircle the DNA containing the mismatch. It is not clear how this protein recognises a mismatch; however it has been proposed that it binds to DNA in a non-specific way bending the DNA in search of mismatches. If a mismatch is found, MutS undergoes conformational changes that kinks the DNA and ultimately unfolds it so that it is recognised by further repair proteins (Wang *et al.*, 2003). The interaction with the mismatch is carried out by a conserved glutamate within the *E. coli* MutS protein (Glu38) which forms a hydrogen bond with the mismatched base (Lebbink *et al.*, 2006).

Following the binding to the mismatch, ATP binds to MutS protein inducing a conformational change. The structural change suffered by MutS helps the binding of another MMR protein, MutL. Similar to MutS, MutL possesses an ATPase activity and it forms homodimers that bind to MutS (Table 1.2). The formation of the MutS-MutL heteroduplex and the presence of ATP in the complex activates the endonuclease activity of another MMR protein called MutH. Monomers of the MutH protein cleave the newly synthesised strand containing the mismatch (Ban and Yang, 1998). The strand discrimination signal that *E. coli* uses to cleave and to remove the DNA mismatch is the unmethylated state of the newly synthesised strand. It has been found that MutH cleaves at the most proximal hemimethylated GATC site, located within

~1 kb of either side of the mismatch. After that, the nicked-strand that contains the mismatch is degraded by a single-strand DNA exonuclease. Depending on the position of the cleavage, either a 3' to 5' or a 5' to 3' exonuclease is used (Cooper *et al.*, 1993). In *E. coli* there are two 3' to 5' exonucleases (ExoI and ExoX) and two 5' to 3' exonucleases (ExoVII and RecJ).

Once the mismatch-containing strand is nicked, it needs to be unwound from the duplex molecule since the exonucleases only act on single stranded DNA. Unwinding is carried out by the UvrD helicase (also known as MutU) starting at the nick and migrating towards the mismatch in the DNA. The recruitment of the UvrD helicase and the unwinding direction is stimulated by MutL, therefore the role of MutL is to stimulate the MutH endonuclease activity and to direct the unwinding and degradation of the mismatch-containing strand (Dao and Modrich, 1998). After the mismatch is removed, the gap created is resynthesised by DNA polymerase III and a DNA ligase seals the nick to complete the mismatch repair process. Single-strand DNA binding protein (SSB) is also involved in the repair process by protecting the template strand after the degradation of the mismatch-containing strand (Modrich and Lahue, 1996).

1.7.2 Mismatch Repair in eukaryotes

Most of the MMR steps used by prokaryotes are conserved in eukaryotes. However, some features such as the dependence of the DNA methylation status to recognise the daughter from the parental strand has only been observed in prokaryotes. In all eukaryotic organisms there are conserved MutS and MutL homologs and they function as heterodimer complexes (Table 1.2). In yeast and humans there are five highly conserved MutS homologs (Msh2 to Msh6). In addition Msh1 is present only in yeast and it is involved in the repair and maintenance of mitochondrial DNA (Reenan and Kolodner, 1992). Furthermore, Msh4 and Msh5 are involved in meiotic recombination but not in DNA repair (their function is discussed in section 1.7.4).

Msh2 has been found to be involved in the recognition of all mismatches in nuclear DNA, whereas Msh6 (also known as GTBP) and Msh3 participate in some distinct types of mismatches. Msh2 forms heterodimers with Msh6 (called MutSα) or with Msh3 (called MutSβ). MutSα and MutSβ have overlapping roles in recognising base mismatches and single base insertion/deletion loops; however MutSβ repairs preferentially insertion/deletion loops (Figure 1.10; Genschel *et al.*, 1998; Umar *et al.*, 1998).

E. coli protein	Homologs Yeast	Humans	General function in yeast and humans
MutS	yMsh1	Not known	Involved in the repair and maintenance of mitochondria DNA
	Msh2-Msh6 (MutSα)		DNA mismatch damage recognition
	Msh2-Msh3 (MutSβ)		Repairs preferentially large insertion-deletion loops but it is partially redundant with MutSα
	Msh4-Msh5		Involved in meiotic recombination
MutL	MutLα yMlh1-yPms1	hMlh1-hPms2	Coordinates events from mismatch binding to DNA repair synthesis
	MutLβ yMlh1-yMlh2	hMlh1-hPms1	≠ Repairs some insertion-deletion loops
	Mlh1-Mlh3 (MutLγ)		Repairs some insertion-deletion loops and participates in meiosis
MutH	Not known		* Nicks nascent unmethylated strand
UvrD	Not known		* Unwinds DNA helix
ExoI, ExoVII, ExoX, RecJ	ExoI		Degrades the mismatch containing strand
SSB	RPA		Protects template DNA from degradation and stimulates excision process
β-Clamp	PCNA		Participates in repair initiation by recruiting MutS to the mismatch and DNA re-synthesis
DNA pol III	DNA pol γ		Involved in repair synthesis
DNA ligase	Ligase I		Seals nicks after repair synthesis

Table 1.2 Function of the mismatch repair proteins in *E. coli* and their homologs in yeast and humans. **E. coli* specific function. ≠Yeast specific function. Yeast Pms1 is equivalent to human Pms2 and yeast Mlh2 to human Pms1. Abbreviations are y, yeast; h, human; SSB, single-stranded binding protein; PCNA, proliferating cell nuclear antigen; RPA, replication protein A.

Four MutL homologs, Mlh1, Mlh3, Pms1 (Mlh2 in yeast), and Pms2 (Pms1 in yeast), have been identified in yeast and humans. Mlh1 forms heterodimers with the three MutL homologs. MutLα is the complex formed by Mlh1-Pms2 (Mlh1-Pms1 in yeast), MutLβ and MutLγ the complexes formed by Mlh1-Pms1 (Mlh1-Mlh2 in yeast) and Mlh1-Mlh3 respectively (Harfe and Jink-Robertson, 2000). The MutLα is involved in the repair of mismatches by interacting with MutSα or MutSβ in yeast and humans. The function of MutLβ in humans is unknown although in yeast is involved in the repair of a subset of insertion-deletion loops. MutLγ also takes part in the repair of some insertion-deletion loops in yeast and humans. For example, mutations in the *MLH2* and *MLH3* genes elevated the mutation frequency at mononucleotide repeat arrays up to 15 times compared to wild type strains in budding yeast, suggesting these two proteins are directly involved in the repair of this type of mismatch. In addition, mutations in the *MLH3* gene have been linked to microsatellite instability in human cells (Harfe *et al.*, 2000; Lipkin *et al.*, 2000).

The recognition of the mismatch by MutSα or MutSβ is facilitated by the proliferating cell nuclear antigen protein (PCNA). PCNA is an ATP dependent protein that can exist as a trimer to form a clamp shape structure that encircles the DNA. PCNA is recruited to the DNA by another sliding-clamp protein, the replication factor C (RFC) (Bowman *et al.*, 2004). It has been found that PCNA interacts with Msh2 and Mlh1 proteins both in yeast and humans (Figure 1.10). Furthermore, a yeast strain containing mutations in the *PCNA* gene displayed increased mutation rates in dinucleotide repeats (Umar *et al.*, 1996; Lee and Alani, 2006). Thus, it is believed that it may transfer or facilitate the binding of MutSα or MutSβ to the DNA mismatch. In addition, studies in yeast have shown that the N-terminal domain of Msh6 and Msh3 interacts with PCNA linking this gene with the mismatch recognition process (Shell *et al.*, 2007).

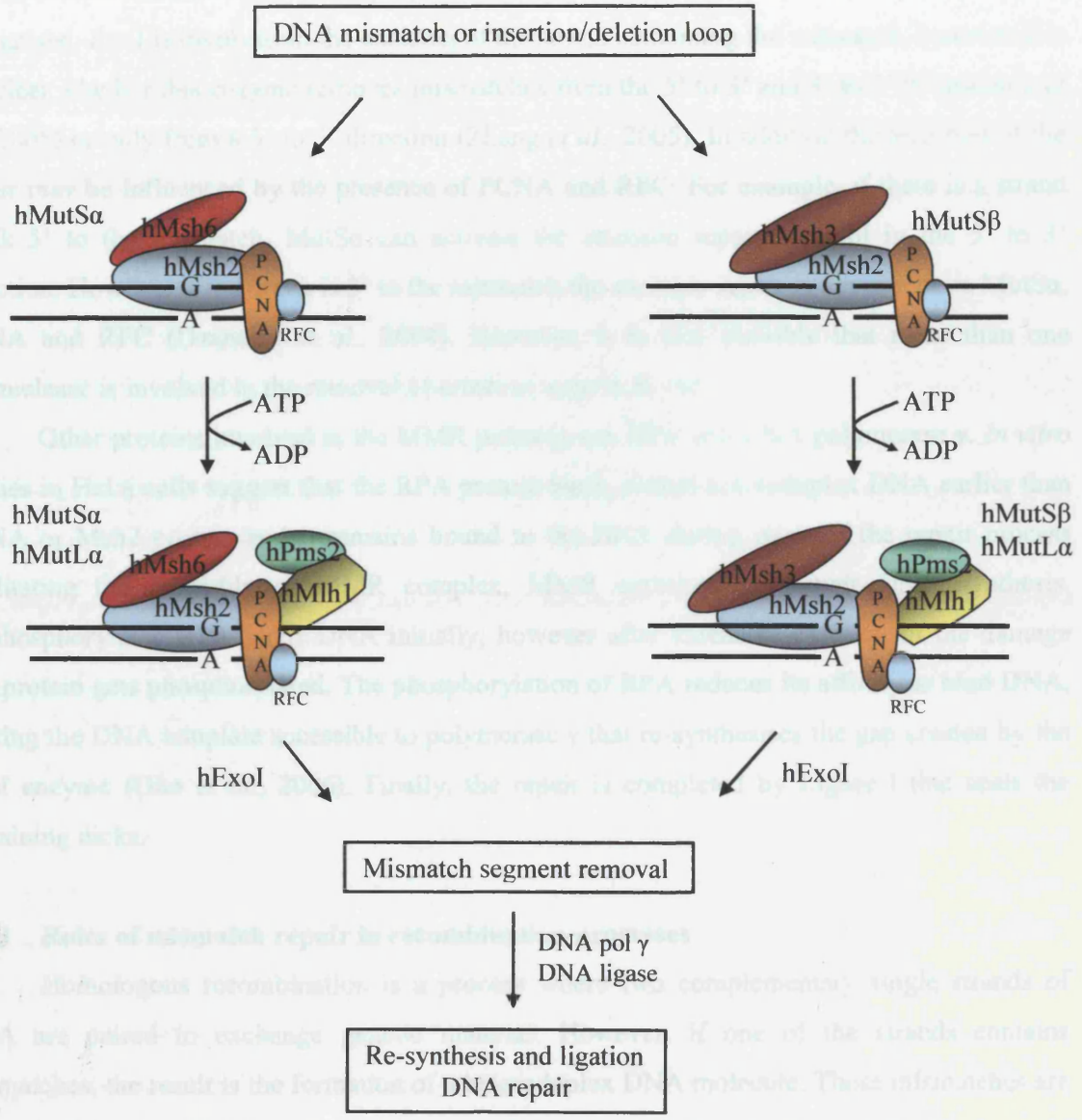


Figure 1.10 Diagram showing the main proteins involved in the mismatch repair pathway. hMutSα and hMutSβ bind to base mispairs and extrahelical loops. PCNA protein may be involved in the mismatch recognition and in the stabilisation of the MMR complex. Once hMutS is bound, hMutLα binds the complex in an ATP dependent way. hExoI removes the mismatch segment while DNA polymerase δ re-synthesise the new DNA fragment.

PCNA protein has also been found to interact with the mismatch excision enzyme ExoI and to control the activity of DNA polymerase γ (Nielsen *et al.*, 2004). Once the mismatch is recognised, ExoI is involved in the excision of the strand containing the mismatch, however it is not clear whether this enzyme removes mismatches from the 5' to 3' and 3' to 5' (Constantin *et al.*, 2005) or only from a 5' to 3' direction (Zhang *et al.*, 2005). In addition the direction of the repair may be influenced by the presence of PCNA and RFC. For example, if there is a strand break 5' to the mismatch, MutSα can activate the excision repair of ExoI in the 5' to 3' direction. However if the break is 3' to the mismatch the excision repair is dependent on MutSα, PCNA and RFC (Dzantiev *et al.*, 2004). However, it is also possible that more than one exonuclease is involved in the removal of errors as seen in *E. coli*.

Other proteins involved in the MMR pathway are RPA and DNA polymerase γ. *In vitro* studies in HeLa cells suggest that the RPA protein binds nicked heteroduplex DNA earlier than PCNA or Msh2 protein and it remains bound to the DNA during most of the repair process facilitating the assembly of MMR complex, MMR excision and repair DNA synthesis. Unphosphorylated RPA binds DNA initially; however after extensive excision of the damage this protein gets phosphorylated. The phosphorylation of RPA reduces its affinity to bind DNA, making the DNA template accessible to polymerase γ that re-synthesises the gap created by the ExoI enzyme (Guo *et al.*, 2006). Finally, the repair is completed by Ligase I that seals the remaining nicks.

1.7.3 Roles of mismatch repair in recombination processes

Homologous recombination is a process where two complementary single strands of DNA are paired to exchange genetic material. However, if one of the strands contains mismatches, the result is the formation of a heteroduplex DNA molecule. Those mismatches are recognised by the MMR pathway and either they are repaired or the recombination event is aborted (reviewed in Harfe and Jink-Robertson, 2000). The repair of mismatches in the heteroduplex molecule results in gene conversion events. However, in *msh2*, *pms1* or *mlh1* mutant yeast strains the mismatches in heteroduplex DNA are not repaired and the mismatch-containing strand segregates (this event is known as postmeiotic segregation) (Alani *et al.*, 1994).

MMR is involved in inhibiting recombination between non-identical sequences (process known as homeologous recombination). This inhibition was first demonstrated by studying intergenic recombination between *E. coli* and *Salmonella typhimurium*. The genomes of these organisms are approximately 20% sequence-divergent and recombination is stimulated dramatically by eliminating *mutS* or *mutL* genes from either *E. coli* or *S. typhimurium* (Rayssiquier *et al.*, 1989). Similar results have been found in *S. cerevisiae* where the length and the homology of the substrate involved in recombination are important for limiting the recombination process, with as little as a single mismatch being enough to inhibit ectopic (non-allelic) recombination (Datta *et al.*, 1997). In addition, *msh2*Δ yeast strains display higher recombinational rates in mitosis and meiosis than *pms1*Δ and *mlh1*Δ strains, suggesting Msh2 protein is one of the main players in suppressing homeologous recombination (Chen and Jinks-Robertson, 1999). In meiosis, MMR proteins are involved in inhibiting cross-over events between diverged sequences. Thus, it has been proposed that the MMR system establishes a genetic barrier to recombination processes between species by restraining exchange of information between diverged sequences (Chambers *et al.*, 1996).

1.7.4 Specificity of mismatch repair proteins in meiosis

The function of some MutS and MutL homologs is limited to meiosis. For example Msh4 and Msh5 proteins lack the ability to recognise mismatches and induce repair, however these proteins are involved in recombination processes during meiosis. Furthermore it has been shown in yeast that the Msh4-Msh5 heterodimer associates with Holliday junctions in meiosis, thus Msh4 and Msh5 proteins may be involved in the stabilisation and/or the resolution of these structures (Ross-Macdonald and Roeder, 1994; Hollingsworth *et al.*, 1995; Snowden *et al.*, 2004). Mutations of either *MSH4* or *MSH5* genes in yeast lead to a reduction in the number of cross-over events and a decrease in spore viability. In addition, homozygous *MSH4*$^{-/-}$ and *MSH5*$^{-/-}$ mice displayed meiosis defects due to chromosomal pairing deficiencies in prophase I leading to meiotic arrest and sterility (Heyer *et al.*, 1999).

Another heterodimer involved in cross-over formation during meiosis is the one formed by Mlh1-Mlh3 proteins (MutLγ). Mutations in either of these genes resulted in a similar phenotype to that seen in *MSH4* or *MSH5* yeast mutants, thus it has been argued that *MLH1*, *MLH3*, *MSH4* and *MSH5* function in a single pathway (Hunter and Borts, 1997; Wang *et al.*, 1999). Although Mlh3 has been found to form a heterodimer with Mlh1 in yeast, evidence in

mouse suggests that Mlh3 may bind to paired chromosomes in prophase I before Mlh1. Furthermore, Mlh3 has been found to associate with other regions of the genome (such as the repetitive centromeric sequences) independently of Mlh1. Thus, it is possible that Mlh3 may function alone or it may form another MutL heterodimer not yet described (Kolas *et al.*, 2005).

1.7.5 Mismatch repair proteins and cancer

Defects in MMR genes have been associated with hereditary non-polyposis colorectal cancer (HNPCC; also known as Lynch syndrome). Patients with HNPCC develop colorectal cancer predominantly in the proximal colon at an early age (~45 years). Apart from colon cancer, families affected with the HNPCC show an increased risk of developing ovary, stomach, endometrium and brain cancers among others (Watson and Lynch, 1993). HNPCC is mainly caused by defects in *MLH1* or *MSH2* gene, however alterations in *MSH6*, *PMS2* and *PMS1* genes have also been found but to a lesser extent (Bronner *et al.*, 1994; Papadopoulos *et al.*, 1994; Miyaki *et al.*, 1997; Hendriks *et al.*, 2006).

Alterations in one of the MMR genes in HNPCC patients leads to a high incidence of instability at simple repeat sequences, a phenomenon called microsatellite instability (MSI) and therefore MSI analysis is one of the main diagnostic markers for HNPCC. In 1998, the National Cancer Institute (USA) set guidelines (known as Bethesda guidelines) for the determination of MSI in colorectal cancer. The Bethesda guidelines suggested that at least five different microsatellites should be used in MSI studies. It is recommended that 2 mononucletides and 3 dinucleotides are analysed. Instability in only one of them is called MSI-L (low) whereas when instability is present in two or more microsatellites it is scored as MSI-H (high) (Boland *et al.*, 1998). MSI is not a unique marker for HNPCC; it is also present in a wide variety of non-colonic tumours such as endometrial, gastric and lung cancers some of them with characterised MMR defects (Fleisher *et al.*, 2001).

Some sporadic colon cancers are also associated with defects in the MMR system. These cancers display similar features as familial cases of HNPCC such as MSI; however promoter methylation not gene mutation is the most common event that leads to inactivation of one of the MMR genes in these patients (Kane *et al.*, 1997). Methylation of *MLH1* promoter is common among sporadic colorectal cancers whereas methylation of the *MSH2* gene is rarely observed. However, methylation of the *MSH2* promoter was recently found in a family that showed inheritance of an HNPCC phenotype. Methylation of *MSH2* was inherited in three successive

generations of this family (Chan *et al.*, 2006) indicating that epigenetic factors can affect MMR gene expression in sporadic and familial colon cancers.

Several studies have been carried out in MMR knockout mice to understand the molecular basis of HNPCC and sporadic colon cancers. Most MMR knockout mice display MSI phenotype and have an increase rate of tumour development. However knockouts for *PMS1*$^{-/-}$ and *MSH6*$^{-/-}$ only showed low instability in mononucletides and dinucleotides respectively. Furthermore, *PMS1*$^{-/-}$ mice did not develop any tumours, suggesting that loss of this gene has a mild effect on genome instability in mouse (Prolla *et al.*, 1998; de Wind *et al.*, 1999). MMR knockout mice (apart from *PMS1*$^{-/-}$) developed different kind of tumours such as lymphomas, gastrointestinal tumours and sarcomas (sarcomas are preferentially developed by *PMS2*$^{-/-}$ knockout mice; Prolla *et al.*, 1998). Surprisingly, none of these mice knockouts developed colorectal cancer, though they displayed genomic instability as seen in humans.

Different cell lines with defects in MMR genes have been established from human colon cancers. Cell lines that are defective for one of the MMR genes are characterised by displaying MSI. However the level of MSI varies according to the MMR defect background. For example, the instability at six microsatellites including di-, tri-, or tetranucleotide repeats was evaluated in the cell line HCT116, which does not express the *MLH1* gene. Instability was present in all the microsatellites analysed, with a high mutation rate (10^{-2} per cell per generation). In contrast, two cell lines (DLD-1 and HCT15) with mutations in *MSH6* only showed very low levels of instability in two out of six microsatellites (only in dinucleotides) (Bhattacharyya *et al.*, 1994). Furthermore, it has been observed that colon cancer cell lines and tumours with MSI are different compared to colon cancers cell lines with stable microsatellites (MSS). For example, MSI colorectal tumours have reduced levels of mutations in genes such as *TP53* as well as showing karyotype stability and low frequency of loss of heterozygosity (LOH) compared to MSS tumours and cell lines (Abdel-Rahman *et al.*, 2001; Gayet *et al.*, 2001; Samowitz *et al.*, 2007). However it is proposed that MMR defects can contribute to the generation of colorectal cancer by promoting mutations in genes that contain microsatellites within their coding regions such as the *APC, K-RAS, BAX* genes (Rampino et al., 1997).

The role of other MMR genes in HNPCC is controversial, for example several missense mutations and truncations of the *EXO1* gene have been found in familial colorectal cancer (Sun *et al.*, 2002). Nevertheless, other studies have revealed that *EXO1* variants are found in normal populations and thus *EXO1* may not be directly implicated in colorectal cancer (Jagmohan-

Changur *et al.*, 2003). Experiments in *EXOI*$^{-/-}$ deficient mice had similar effects as mutations in other MMR genes, for example, there is an increase in the probability to develop lymphomas and also instability at simple tandem repeats is observed (Wei *et al.*, 2003). Thus it is likely that *EXOI* mutations in humans have similar effects as mutations in other MMR genes.

1.7.6 DNA damage response and its relationship with mismatch repair

MMR proteins are able to recognise certain types of DNA damage, for example damage caused by alkylating agents, UV light and reactive oxygen species. In addition, MMR proteins of mammalian cells can activate cell cycle checkpoints and apoptosis after recognising DNA damage (Karran, 2001). The role of MMR proteins in inducing apoptosis after certain types of DNA damage was observed for the first time in *mutS* and *mutL E. coli* mutant strains. These mutant strains were not killed by low doses of methylating agents such as N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) or N-methyl-nitrosourea (MNU). This effect has also been observed in human cancer cells with mutations in MMR genes; however, resistance to alkylating agents has not been observed in *S. cerevisiae* following deletion of a MMR gene (Umar *et al.*, 1997).

Methylating agents are cytotoxic and can damage DNA by methylating guanine or thymine residues in DNA. However, cells have evolved mechanisms to correct such damage by, for example, activation of the DNA methyltransferase (MGMT) that repairs DNA O^6-methylguanines (O^6-meGua) or O^4-methylthymine aberrations. However, if MGMT does not repair this damage, the MMR machinery lead by MutSα and MutLα induces cell death (Duckett *et al.*, 1996). During replication, insertion of a base opposite to a methylated guanine or thymine initiates the process of repair by MMR. However, when the damage is located in the template strand, repetitive futile cycles of excision and resynthesis occur that lead to cell death. The mechanisms that lead to apoptosis in cells exposed to MNNG are not clear and may involve more than one pathway, however the p53 protein is phosphorylated in response to DNA damage signals and it is dependent on MutSα and MutLα. This suggests that p53 is a key component of cell-cycle arrest or apoptosis upon failure to repair the DNA damage by MMR. In addition, Pms2 and Mlh1 have been found to interact with p53, implying a strong relationship between MMR and the role of p53 in regulating cell growth (Duckett *et al.*, 1999; Chen and Sadowski, 2005). Furthermore, it has been found that Msh2 and Mlh1 interact with Atr and Atm kinases respectively. These interactions are enhanced by the presence of methylating agents or UV light and in the case of Msh2, its interaction with Atr is required to phosphorylate Chk1.

Phosphorylation of Chk1 is required for the activation of S-phase checkpoint to inhibit DNA synthesis in response to damage agents such as MNNG (Brown *et al.*, 2003; Wang and Qin, 2003).

1.8 DNA repair proteins at telomeres

DNA repair pathways are activated in response to DNA damage and therefore they help to maintain chromosome integrity. However, at telomeres some of the repair pathways can have deleterious effects and that is the case of the NHEJ pathway.

The NHEJ pathway is involved in the generation of circular or dicentric chromosomes when telomeres are too short or when Trf2 protein is inhibited or depleted from cells (van Steensel *et al.*, 1998; Smogorzewska *et al.*, 2002). The joining of two telomeres is mediated by ligase IV which requires the degradation of the 3' overhang in order to join two telomeres together. The NER complex formed by Ercc1 and Xpf is the main candidate to achieve this degradation (Figure 1.11). Cells deficient in the Ercc1/Xpf complex resulted in the retention of the overhang after Trf2 depletion, suggesting that Trf2 protein protects telomeres from the endonuclease activity of the Ercc1/Xpf complex (Zhu *et al.*, 2003). However, apart from degrading the 3' overhang, this complex is involved in preventing recombination events between telomeres and interstitial telomeric sequences. *ERCC1*$^{-/-}$ mouse cells contain numerous telomeric DNA double minute chromosomes; which are believed to be generated by interstitial recombination events (Zhu *et al.*, 2003). Thus, the presence of Ercc1/Xpf proteins may have a dual role on telomeres; on one side they protect and repair them however under certain conditions they can also generate further instability.

Other members of the NHEJ pathway such as the Ku complex (Ku70 and Ku80) and DNA-PKcs are found at telomeres and they contribute to their protection (d'Adda di Fagagna *et al.*, 2001). For example, inactivation of the Ku complex in *S. cerevisiae* and *S. pombe* leads to telomere shortening and end protection defects (Baumann and Cech, 2000). However, loss of Ku not only led to telomere shortening by the elimination of the 3' overhang protection but also by the inhibition of telomerase recruitment to telomeres (Bertuch and Lundblad, 2003). Similar effects have been found in mouse embryo fibroblasts from *KU70*$^{-/-}$ and *KU80*$^{-/-}$ knockout mice, however telomere length was not affected in mouse cells with defects in other members of the NHEJ pathway such as ligase IV and Xrcc4 (d'Adda di Fagagna *et al.*, 2001).

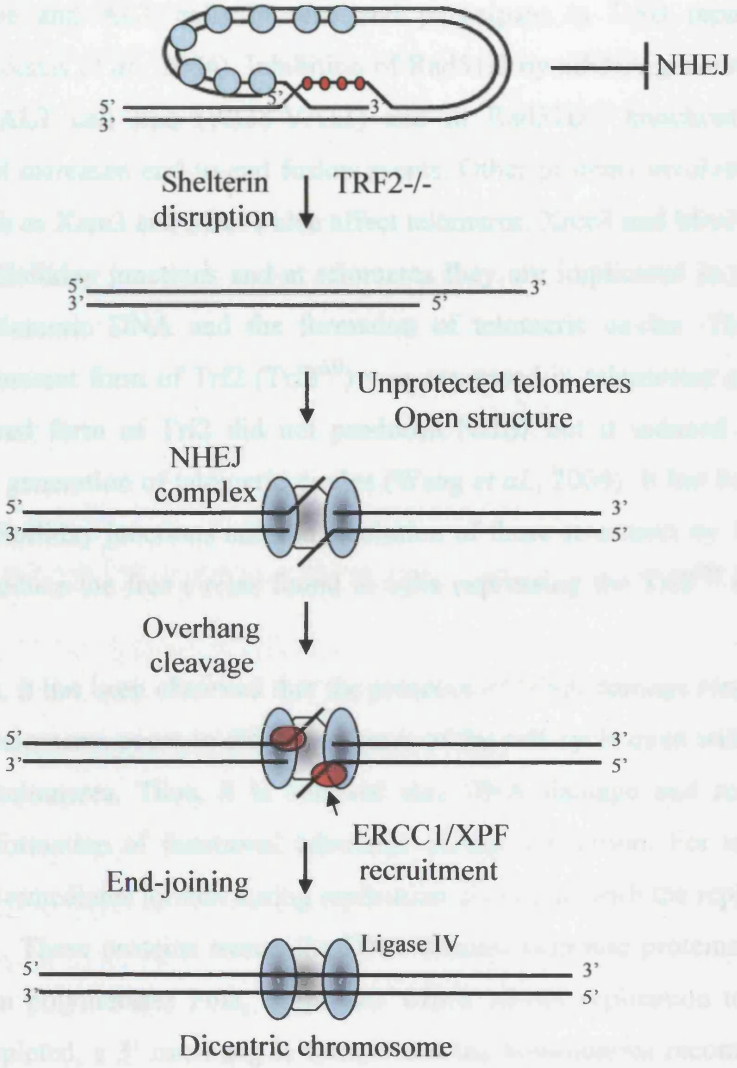


Figure 1.11 Mechanism of NHEJ at telomeres. Loss of telomere protection by disruption of shelterin or telomere shortening would produce an open telomere structure. Under these circumstances, the NHEJ complex may recruit the Ercc1/Xpf complex which degrades the 3' overhang of the telomeres. Ligase IV may then join the two telomeres together creating a dicentric chromosome. Figure adapted from de Lange (2005).

Apart from the proteins involved in NHEJ, many of the repair proteins are also associated with telomeres and have a direct impact in their function. For example, proteins involved in homologous recombination, such as Rad51D, localise to mammalian telomeres in telomerase positive and ALT cells *in vivo* and participate in DSB repair and telomere maintenance (Tarsounas *et al.*, 2004). Inhibition of Rad51D by siRNA generated shortening of telomeres in an ALT cell line (WI38-VA13) and in Rad51D$^{-/-}$ knockout mouse embryo fibroblast cells and increased end-to-end fusion events. Other proteins involved in homologous recombination such as Xrcc3 and Mre11 also affect telomeres. Xrcc3 and Mre11 are involved in the resolution of Holliday junctions and at telomeres they are implicated in the generation of sudden loss of telomeric DNA and the formation of telomeric circles. These events were observed when a mutant form of Trf2 (Trf2$^{\Delta B}$) was expressed in telomerase positive and ALT cells. This truncated form of Trf2 did not produced NHEJ but it induced sudden telomere truncation and the generation of telomeric circles (Wang *et al.*, 2004). It has been proposed that t-loops can form Holliday junctions and the resolution of these structures by Xrcc3 and Mre11 proteins would produce the free circles found in cells expressing the Trf2$^{\Delta B}$ allele and seen in ALT cells.

In addition, it has been observed that the presence of DNA damage response signals and repair proteins at telomeres occur in different phases of the cell cycle even without the presence of dysfunctional telomeres. Thus, it is believed that DNA damage and repair proteins are essential for the formation of functional telomeres during replication. For instance, telomere single-stranded intermediates formed during replication are coated with the replication protein A (RPA) and Mre11. These proteins recruit the DNA damage response proteins Atm/Atr, Rad17 and the replication polymerases Polα, Polβ/Fen1 which allows replication to continue. After replication is completed, a 3' overhang is formed and the homologous recombination proteins Rad51/Rad52/Xrcc3 localise to the end of the telomere presumably to promote the invasion of the 3' single-stranded overhang to the double-stranded DNA to form a D-loop structure (Verdun and Karlseder, 2006).

1.9 Mismatch repair proteins and their role in telomere stability

As mention above, defects in the MMR pathway is one of the key players in cancer formation in HNPCC patients. Patients with HNPCC show increase mutation frequency mainly at DNA repeats such as microsatellites (MSI). Thus, human telomeres composed of arrays of

TTAGGG repeats are likely to suffer mutations in MMR deficient cells. Mutation analysis at telomeres has been carried out in colon cancers that display MSI and it was found that telomere mutations exist in some MMR defective tumours and are probably generated by intra-telomeric events (Pickett *et al.*, 2004). Furthermore, the mutation frequency in a randomly collected panel of sporadic colon cancer rose from 5.8% to 35% per telomere in colon cancers with characterised mutations in the *MSH2* gene. These results show that defects in the MMR pathway not only affect microsatellite instability but also telomeres *in vivo*. However the biological significance of telomere instability in MMR deficient cells is not known. Previous experiments in yeast have shown that mutations in one MMR gene (*MSH2*, *MLH1* or *PMS1*) in telomerase deficient strains alleviate the barrier of homeologous recombination leading to the survival of those strains *via* a recombination-base mechanism (Rizki and Lundblad, 2001). Thus, MMR proteins are important in limiting the mutation and recombination events at telomeres.

1.10 Analysis of single human telomeres

The analysis of chromosome-specific human telomeres has proved to be difficult because of the high content of repeat sequences and segmental duplications in the subtelomeric regions (Riethman *et al.*, 2004). However, sequence analysis of human subtelomeres at some chromosome ends has revealed chromosome-specific sequences near the start of the telomere that facilitate the analysis of individual telomeres. Furthermore, some chromosome ends contain regions of sequence variation between individuals as seen by a high density of single nucleotide polymorphisms (SNPs) adjacent to some telomeres. At the Xp/Yp and 12q subtelomeres, the SNPs form regions of association (linkage disequilibrium) that form haplotypes. Thus it has been possible to analyse single telomeres by designing allele-specific primers that amplify telomeres from the adjacent chromosome specific sequence (Baird *et al.*, 1995, 2000). Chromosome-specific telomere analysis has also been carried out at other ends such as the 2p, 11q, 16p/q and 17p (Coleman *et al.*, 1999; Britt-Compton *et al.*, 2006).

1.10.1 Organisation of the Xp/Yp, 12q and 16p/16q telomere-adjacent region

The terminal 1 kb of the human pseudoautosomal region (PAR1) contains a short interspersed nuclear element (SINE) and a minisatellite (Royle., 1995). Sequence analysis of the 850 bp adjacent to the Xp/Yp telomere has revealed a high level of DNA sequence variation. Thirteen polymorphic bases (SNPs) have been identified in Caucasian samples while seventeen

have been detected in African DNAs. In addition a 10 base pair segment forms an insertion/deletion polymorphism (in/del) close to the start of the telomere. There is strong linkage disequilibrium across those polymorphic sites resulting in a limited number of haplotypes in populations (3 haplotypes in Caucasians and 4 in Africans; Baird *et al.*, 1995). Similar to the Xp/Yp telomere, the 12q telomere-adjacent region exhibits a high level of polymorphism and strong linkage disequilibrium. Sequence analysis of the 1,870 bp adjacent to the 12q telomere has identified twenty two SNPs and a 30 bp in/del polymorphism. Twenty of the SNPs are in almost complete linkage disequilibrium, generating three haplotypes which account for 80% of the 12q telomeres in Caucasians. A subset of the population (6%) contains a 1,439 bp deletion in the flanking telomere region (12qΔ allele). The remaining 13% of 12q telomeres do not amplify with 12q allele-specific primers and were considered as null alleles (Baird *et al.*, 2000).

Telomeres on chromosome 16 are polymorphic for the presence or absence of sequences immediately internal to the telomeres. A subset of chromosome 16 (6%) contains a sequence that can be present adjacent to the 16p or 16q telomere. When present in an individual, it is possible to use this sequence to amplify the associated 16p/16q telomere (Coleman *et al.*, 1999).

1.10.2 Mapping the distribution of telomeres and sequence variant repeats by Telomere Variant Repeat-PCR (TVR-PCR) method

Human telomeres are formed by arrays of TTAGGG repeat units; however the proximal region of telomeres contains variant repeat units such as TTGGGG, TCAGGG, TGAGGG (Allshire *et al.*, 1989) and CTAGGG (Hills unpublished) among others. These telomere repeats are not influenced by the activity of telomerase and are commonly found within the proximal 2 kb of the telomeric DNA while the rest of a telomere is formed by the consensus TTAGGG repeat. In order to study allelic variation at the proximal end of telomeres, telomere variant repeat by PCR (TVR-PCR) was developed (Baird *et al.*, 1995). This technique is based on the amplification of telomeres using radioactively labelled allele-specific primers which anneal to one of the SNPs located in the telomere-adjacent region (or to polymorphic regions on some chromosomes) together with a primer that anneals to one type of telomeric repeat (Figure 1.12). The PCR products are resolved on polyacrylamide gels as a 6 bp ladder. Thus a telomere code of the distribution of sequence variant and TTAGGG repeats can be determined. Using the TVR-PCR method, it has been possible to study the stability and molecular dynamics of single

telomeres, showing that telomeres are highly polymorphic as a consequence of a high mutation rate (Baird *et al.*, 1995; Baird *et al.*, 2000; Varley *et al.*, 2002).

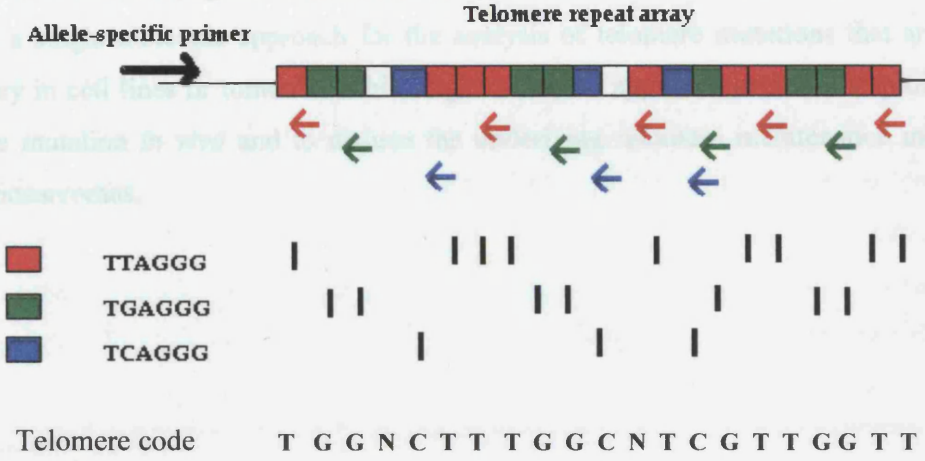


Figure 1.12 Telomere maps generated by TVR-PCR. Allele-specific amplification is carried out using a radioactively labelled primer that anneals to one of the SNPs in the telomere-adjacent sequence in a heterozygous individual together with TAG-TELW (anneals to TTAGGG repeats) or TAG-TELX (anneals to TGAGGG repeats) or TAG-TELY (anneals to TCAGGG repeats) primers. Following size resolution, telomere maps can be read into codes. T (TTAGGG), G (TGAGGG), C (TCAGGG) and N (non-amplifying repeats).

1.11 Aims of project

It has been shown that the repetitive conformation of human telomeres predisposes them to instability and this instability is elevated in colon cancers with defects in MMR (Pickett *et al.*, 2004). The aim of this project was to study the role of MMR on telomere stability by examining the telomere mutation frequency in cells with defects in the MMR system and by disrupting expression of the *MSH2* gene in a normal fibroblast cell line. Another aim of this project was to develop a single molecule approach for the analysis of telomere mutations that arise at a low frequency in cell lines or tumours. This single molecule approach was then exploited to study telomere mutation *in vivo* and to deduce the underlying telomere maintenance mechanism in some liposarcomas.

Chapter 2

Materials and Methods

2.1 Materials

2.1.1 Chemicals, molecular biology reagents and equipment

Chemicals were mainly obtained from Fisher Scientific (Loughborough, UK), Geneflow (Stafford, UK), FMC Bioproducts (Rochland, USA), Sigma-Aldrich Company (Poole, UK), Fisons (Loughborough, UK) and Serva (Heidelberg, Germany).

Molecular biology reagents were obtained from Advanced Biotechnologies (Leatherland, UK), Bioline (London, UK), Boehringer Mannheim (Lewes, UK), Life Technologies (Paisley, UK), National Diagnostics (Hull, UK), New England Biolabs (Hitchin, UK), Pharmacia (Milton Keynes, UK) and Roche Molecular Biochemiclas (USA). Radiochemicals were obtained from PerkinElmer (Austria) and tissue culture reagents from Gibco (Paisley, UK).

Most of the equipment used to perform this work was obtained from Amaxa Biosystems (Koel, Germany), Bio-Rad (Hemel Hempstead, UK), Eppendonf (USA), GE healthcare (USA), Genetic Research Instrumentation (Dunmow, UK), Hybaid (Teddington, UK) and Perkin-Elmer/Applied Biosystems (Beaconsfield, UK).

2.1.2 Oligonucleotides

DNA oligonuleotides were synthesised by Sigma-Aldrich Company (Poole, UK). siRNA were synthesised and annealed by Ambion (UK). The oligonucleotide sequences are shown in Appendix I.

2.1.3 Cell lines, tumour samples and DNA plasmids

Colon cancer cell lines (HCT15, LoVo, LS411, Vaco5 and LS180) were provided by Dr. Ian Tomlinson (Cancer Research UK, London). CCD34-Lu primary cell line was obtained from the American Type Culture Collection. DNAs from lymphoblastoid cell lines were obtained from the Centre d'Etude du Polymorphism Humain (CEPH, Paris, France).

A panel of normal colon and colon carcinoma samples from patients with sporadic colorectal cancer was collected previously at the Leicester Royal Infirmary (Pickett *et al.*,

2004). Liposarcoma DNA samples were provided by Dr. Nadia Zaffaroni (Istituto Nazionale Tumori Milan, Italy) and are described in Costa *et al.*, 2006.

pSuperior.puro plasmid was a gift from Dr. Cristina Tufarelli (University of Leicester); the pCIneo-hTERT plasmid was provided by Dr. Roger Reddel (The Children's Medical Research Institute, Australia); pmaxGFP plasmid was obtained from Amaxa Biosystems (Cologne, Germany).

2.2 Methods

2.2.1 Tissue Culture

Tissue culture was performed in a class II laminar flow hood in a designated tissue culture area. Cells were grown in 5% CO_2 with high humidity at 37°C in an incubator (Sanyo). HT1080 and colon cancer cell lines were grown in Dulbecco's D-MEM with Glutamax-I (Gibco, Paisley, UK) and 10% foetal calf serum. CCD34-Lu cell line was grown in Dulbecco's MEM with Glutamax-I, 20% foetal calf serum and 1x non-essential amino acids (Gibco, Paisley, UK).

2.2.1.1 Cloning of tissue culture cells

Cloning experiments were performed as follow: cells were grown in 25 cm^2 flasks until they reached confluence. Trypsin was added to the flask for 5 minutes in order to detach the cells and to form a single cell suspension. The trypsin reaction was stopped by adding cell culture medium supplemented with foetal calf serum. In all cloning experiments, cells were centrifuged and the pellet was re-suspended in 10 ml of media containing 20% of Gold serum (PAA laboratories, Austria). The cells were counted and serial dilutions were made to give a concentration of 8 cells/ml. Overall, 100 µl of the cell suspension was pipetted into each well of a 96-well plate and approximately 8 plates were seeded in this way for each cell line.

The plates were observed periodically in order to determine the wells that contained single colonies. Once the cells covered almost all the base of the well, the medium was removed and the cells were washed using 100 µl of 1 X PBS (137 mM sodium chloride, 2.7 mM potassium chloride, 4.3 mM disodium phosphate, 1 mM potassium dihydrogen phosphate; MP biomedicals, Germany). The PBS was removed and 100 µl of trypsin was added per well until cells formed a single cell suspension. The cells were transferred to a 25 cm^2 flask containing 8.5 ml of media and grown until they were harvested for DNA extraction.

2.2.2 pCIneo-hTERT plasmid preparation

In order to have sufficient amount of pCIneo-hTERT plasmid (containing the hTERT cDNA) to performed electroporation experiments of human cells, 20 ng of the pCIneo-hTERT plasmid were electroporated into XL-1 blue *E. coli* competent cells (genotype; *endA1 gyrA96 thi-1 recA1 relA1 supE44 hsdR17 lac* [F' Tn10 proAB $lacI^q$ Δ(lacZ)M15 (Tet^R)]; Stratagene U.S.) as described in Sambrook and Russell, 2001. Colonies were selected in ampicillin and recovered by Mini-preps (Quiagen). The plasmid was linearised using *Bgl*II, purified by phenol/chloroform extraction and the DNA was ethanol precipitated. The DNA pellet was resuspended in sterile-distilled water.

2.2.2.1 Electroporation of human cells

The human cells used for electroporation were trypsinised, added to medium containing serum, centrifuged and washed twice in MEM medium (Gibco, Paisley, UK). The cells were adjusted to a final concentration of 2.5 X 10^7 cells/ml and 1 X 10^7 cells (400 µl) were used per electroporation. The cells to be electroporated (resuspended in MEM medium) were transferred to a cuvette and 6 µg of plasmid DNA (10 - 15 µl) was added to the cell suspension. The cells were electroporated using a Gene Pulser XCell device (BioRad) with 250 V and constant field of 250 µF. After electroporation, ~1.5 X 10^6 cells were seeded per 25 cm^2 flasks. Antibiotics were added 24 hrs (G-418) or 48 hrs (Puromicyn) after electroporation. The media was changed every 72 hrs after the antibiotics were added.

2.2.3 Stable RNA interference experiments

2.2.3.1 pSuperior.puro plasmid preparation

The oligos that contained the shRNA-expressing sequence were diluted in distilled water to get a final concentration of 3 µg/ml and annealed as follow: 1 µl of each oligonuclotide was mixed in 48 µl of annealing buffer (10 mM Tris pH 7.5, 50 mM NaCl, 1 mM EDTA) and incubated at 90°C for 5 minutes. The mixture was allowed to cool down to room temperature. Once annealed, the oligos were ligated into pSuperior.puro vector.

The pSuperior.puro vector was linearised with *Bgl*II and *Hind*III restriction enzymes (New England BioLabs) and the linear vector was gel purified on a 1% LE agarose using the MinElute Gel Extraction Kit (Qiagen). The annealed oligos (2 µl) were ligated into 300 ng of linear pSuperior.puro vector using 400 units of T4 ligase in a 10 µl final volume. The reaction was incubated overnight at room temperature. The plasmid was ethanol precipitated and the pellet resuspended in water. The plasmid was transformed by electroporation into

XL-1 blue *E. coli* competent cells as described before. Prior to transfection into mammalian cells the insert was sequenced (Appendix III), the pSuperior.puro plasmid was linearised with *Hind*III restriction enzyme and purified by phenol/chloroform extraction.

2.2.3.2 Nucleofection

Transfection of pSuperior.puro into human cells was performed by nucloefection, which delivers the plasmid directly into the nucleus. Nucleofection was performed using a nucleofector device (Amaxa Biosystems, Koel, Germany) with settings and solutions provided by the manufacturers. Prior to transfection, the cells were grown until they reached ~80% confluence, the medium was removed and the cells were washed once with PBS. The cells were harvested and counted to determine cell density using a haemocytometer. The cells were centrifugated at 200xg (900 rpm) in a bench centrifuge (Sorvall Legend, Labcare, UK) for 10 min in warm medium. The cell pellet was resuspended to give a final concentration of 2 X 10^6 cells in 100 µl of Nucleofector solution (Amaxa Biosystems) per experiment. Resuspended cells were mixed with 5 µg of plasmid DNA (pSuperior.puro) or 2 µg pmaxGFP (control plasmid). Cells were placed in a cuvette and transfection was carried out using Nucleofector programs set in the device. The programs were chosen according to the best transfection efficiency and survival per cell type (described in next section). After transfection, the cells were removed from the cuvette and incubated at 37°C in 500 µl RPMI medium for 10 min and then 5 X 10^5 cells were transferred to 6-well plates using 2 mls of pre-warmed medium. Puromycin selection was added to the cells 48 hrs after transfection.

2.2.3.3 Visualisation of GFP-transfected cells

In order to obtain the best transfection efficiency and survival of human cells prior to nucleofection, the cells were transfected with pmaxGFP control plasmid (Amaxa Biosystems, Germany) using different programs set in the Nucleofector device. The cells were prepared as described above and were seeded on coverslips in 6-well plates with 2 mls of the appropriate medium. Two days after transfection the medium was removed and the cells were washed with PBS. The coverslips were transferred to clean plates and the cells were fixed using cold methanol at -20°C for 20 min or 3.7% paraformaldehyde for 10 min at room temperature. After that incubation time, the cells were washed twice with PBS and the nuclei were stained with DAPI (Boehringer-Mannheim, Germany) at a concentration of 50 ng/ml for 5 minutes. Visualisation of DAPI and GFP was carried out in a fluorescence

microscope (Axioskop 2, Zeiss, Germany) using a digital camera (Photometrix Quantix) and the Smart Capture software.

2.2.4 Delivery of siRNA by electroporation for transient downregulation

Electroporation was used in order to deliver annealed siRNA into human cell lines for transient experiments. The cells were grown until they reach approximately 80% confluence and ~8 X 10^6 cells were used per electroporation. Prior to electroporation the cells were washed and resuspended in 100 µl Opti-MEM medium (Invitrogen) and either 300 or 600 pmols of dsRNA oligonucleotides were mixed with the resuspended cells and placed on ice for 5 min in a 2mm gap-cuvette. The conditions used for the electroporation were: 160 V, 500 µF (program 3T3, Gene Pulser XCell, BioRad). After electroporation, the cuvette that contains the cells was incubated on ice for 5 min and approximately 5 X 10^5 cells were plated on a 25 cm^2 flask. The cells were detached from the flasks by trypsinization and harvested for Western blot analysis by centrifugating the cells at 1100 rpm in a bench centrifuge (Sorvall Legend, Labcare, UK) after 46 and 72 hrs of transfection.

2.2.5 DNA extraction from tissue culture cells

Cells were detached by trypsinisation and added to medium containing 10% foetal calf serum to stop the trypsin degrading the cells. The cells were centrifugated and the pellet resuspended and washed twice in PBS. The final cell pellet was resuspended in 250 µl (for ~ 1 X 10^6 cells) of 1 X SSC (15 mM sodium citrate, 150 mM sodium chloride) buffer. Cells were lysed by adding 250 µl lysis buffer (100 mM Tris-HCl pH7.5, 100 mM NaCl, 10 mM EDTA, 1% sarkosyl). RNase (final concentration 10 mg/ml) was added to the mixture for 20 minutes in order to degrade RNA. Proteins were digested by adding proteinase K at a final concentration of 20 mg/ml and incubated in a water bath at 55°C for 5 to 6 hrs. DNA was extracted from the solution with an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) by gently mixing to form an emulsion. The organic and aqueous phases were separated by centrifugation at 13000 rpm (Eppendorf 5415 centrifuge) for 6 min at room temperature using Phase-lock gel tubes (Eppendorf). The aqueous phase was removed and the DNA was precipitated using 0.1 volume of 2 M sodium acetate (pH 5.6) and 2.5 volumes of 100% ethanol. The DNA was resuspended in 100 to 500 µl of 1X TE (10mM Tris, 1mM EDTA). The concentration of the DNA was estimated using a spectrophotometer (Eppendorf biophotometer).

2.2.6 Protein extraction from tissue culture cells

The cells were treated as described in the DNA extraction method, however the cell pellet was not resuspended in 1X SSC solution, instead it was frozen with dry ice and kept at -80°C until proteins were extracted. Protein extraction was performed by adding 30 µl of lysis buffer (50 mM Tris, 150 mM NaCl, 1 mM EDTA, Triton 1%) and 1:10 dilution of protease inhibitor cocktail (Sigma) for ~1 X 10^6 cells. The mixture was incubated at 4 °C for 20 min and centrifugated for 10 min at 13000 rpm (Eppendorf 5415 centrifuge) at 4 °C. The supernatant containing total protein extract was transferred to a new tube and stored at -80°C.

Protein concentration was determined using the Bradford method (Sambrook and Russell, 2001) by the addition of 1 ml Coomassie blue staining solution (for 500 mls of solution: 50 mg Brilliant Blue G dye dissolved in 25 ml of 95% ethanol and 50 mls orthophosphoric acid) to ~ 3 µl of protein extract. The solution was incubated for 5 min at room temperature and the protein concentration was determined from BSA concentration standards read at OD_{595} in a spectrophotometer (Eppendorf biophotometer).

2.2.7 DNA amplification

DNA amplification was carried out using the polymerase chain reaction (PCR) technique. PCR was performed using from 50 to 100 ng of genomic DNA, 0.5 to 1 µM of each primer, 0.1 U/µl Taq polymerase (Abgene) and 1X PCR buffer as described in Jeffreys *et al.* (1988) and kept as an 11X stock: 45 mM Tris-HCl pH8.8, 11 mM $(NH_4)_2SO_4$, 4.5 mM $MgCl_2$, 6.7 mM β-mercaptoethanol, 4.4 mM EDTA, 1 mM dATP, 1 mM dCTP, 1 mM dGTP, 1 mM dTTP and 113 µg/ml BSA. PCRs were performed in 10 µl final volume. PCR reactions were carried out using a PTC200 DNA Engine thermal cycler (MJ Research, USA) and the reactions were set up in a category II laminar flow hood to avoid contamination.

2.2.7.1 Amplification of Bat-26 microsatellite

Small-pool PCR was carried out to amplify Bat-26 microsatellite by diluting the DNA to ~30 pg per PCR reaction. PCR was carried out as described previously but the final primer concentration (Bat26-F and Bat26-R) was reduced from 0.1 µM to 0.02 µM and 0.15 µCi of [α-^{32}P] dCTP was added to the reaction. PCR conditions were as follows: initial incubation at 95°C for 5 min and 40 cycles of 95°C for 30 sec, 58°C for 40 sec, 70°C for 1 min and a final extension of 70°C for 10 min. The products generated (121 bp) were resolved on 8% polyacrylamide denaturing gels for 2 hrs at constant 55 W. The gel was

dried and the radioactive phosphorous detected using a phosphorImager device (Typhoon 9410, GE healthcare).

2.2.7.2 Bacteria colony PCR

In order to identify transformed *E. coli* after electroporation with plasmids, colony PCR was carried out. Colony PCR was conducted in a 10 µl reaction using 1X PCR buffer (Jeffreys *et al.*, 1988), 0.1 µM of pSup-1 and pSup-2 primers and *E. coli* taken directly from Petri dishes with a pipette tip. The reactions were cycled under the following conditions: initial incubation step of 96°C for 5 min and 28 cycles of 96°C for 20 sec, 65°C for 30 sec and 70°C for 1 min.

2.2.7.3 Single telomere length analysis (STELA)

STELA technique was used to amplify telomeres and further details of the technique are described in Chapter 4. The DNA used for STELA was digested with *EcoR*I restriction enzyme, quantified using a spectrophotometer (Eppendorf biophotometer) and diluted to 10 ng/µl (stock dilution). To obtain the final DNA concentration (30 – 250 pg/µl), the DNA was diluted in dH_2O containing tRNA (1 ng/µl) as a carrier. A linker oligonucleotide (telorette) was added to the DNA dilution to get a 0.02 µM final concentration. STELA was carried out in a 10 to 20 µl reaction containing 30 – 250 pg/µl of DNA with telorette, 1x PCR buffer, 0.5 µM of telomere or allele specific and teltail primer and 1 unit of a mixture of *Taq* (ABgene) and *Pwo* (Roche) polymerase (10:1). The reactions were cycled as follow: 25 cycles of 96°C for 20 sec, 66°C (XpYp E2, 12q-STELA, CB38A-F or Nitu14eC) or 67°C (allele-specific primers XpYp-427G/415C) for 30 sec and 68°C for 10 min. Products were size separated in 0.8% HGT agarose using 20 X 40 cm gels.

2.2.8 Agarose gel electrophoresis

Agarose gel electrophoresis was carried out using 0.5X TBE (1X TBE, 45 mM Tris-borate, 1 mM EDTA), containing 0.5 µg/ml ethidium bromide. The concentration of agarose varied according to the DNA fragment to be resolved. 1% LE agarose (supplied by SeaKem) was used in most cases but HGT agarose (SeaKem) was used at lower concentrations (up to 0.7%) to separate fragments of more than 10 kb. DNA was visualised using UV transilluminator (UVP High Performance transilluminator, UVP Life Sciences, Cambridge, UK).

2.2.9 Telomere Variant Repeat Mapping by PCR (TVR-PCR)

TVR-PCR was carried out in three parallel reactions, each containing a different telomere primer together with a radioactively labelled allele-specific primer (Baird *et al.*, 1995). The allele-specific primer was 5′ end-labelled at room temperature overnight. The labelling reaction (2 µl) contained 0.5 µM primer, 1x polynucleotide kinase buffer (70 mM Tris-HCl, 10 mM $MgCl_2$, 5 mM dithiothreitol), 2 units of T4 polynucleotide kinase and 0.13 µCi [γ^{32}P] ATP. The labelled primer (2µl) was added to a PCR reaction containing 0.4 µM of one of the telomere repeat primers (TAG-TelW, TAG-TelX and TAG-TelY; Baird *et al.*, 1995), 1 unit of Taq DNA polymerase, 100 ng of genomic DNA and 1X PCR buffer in a total volume of 10 µl. PCR reactions were cycled 21 times under the following conditions: 96°C for 20 seconds, 67-67.5°C for 30 seconds (according to the allele-specific primer used) and 70°C for 3 minutes.

At the end of the PCR, 5 µl of formamide loading dye (95% formamide, 10 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue) were added to each reaction. The samples were denatured at 100°C for 3 min before they were loaded in a denaturing polyacrylamide gel containing 5% (for 12q telomere) or 6% (for Xp/Yp, 16p/q and CB54 telomeres) acrylamide (19:1 acrylamide:bisacrylamide, Geneflow, UK), 7.67 M urea and 1X TBE. The acrylamide gel was electrophoresed at 100 W (~50°C) for 4.5 to 5 hours (12q telomere) or for 3.5 hrs (Xp/Yp, 16p/q and CB54 telomeres). The gel was transferred on 3MM paper (Whatman) and dried for 2 hours at 80°C. The dried gel was exposed overnight to storage phosphor screens (GE healthcare) and the signal detected using a PhosporImager (Typhoon 9410, GE healthcare).

2.2.10 Southern blotting

Southern blotting was used to test for the presence of a specific DNA sequence by using agarose gel electrophoresis to size separate PCR or genomic DNA. The products were transferred to a nylon membrane for probe hybridisation in order to detect the desired DNA fragment.

The agarose gel was incubated in an acid depurinating solution (0.25 M HCl) for 10 min (2X 5 min), alkali-denaturing solution (0.5 M NaOH, 1.5 M NaCl) for 30 min (2X 15 min) and neutralising solution (0.5 M Tris-HCl, 3 M NaCl) for 30 min (2X 15 min). The DNA was transferred to a nylon membrane (MagnaProbe, GE Osmonics) previously soaked in 2X SSC by capillary transfer using 20X SSC as transfer buffer. In most cases, the transfer was carried out overnight; the membrane was dried in an 80°C oven for 15 min and cross-

linked by exposing the blot to 7 x 10^4 J $(cm^2)^{-1}$ of UV light in an Ultraviolet Crosslinker (Amersham Biotechnology).

2.2.10.1 Probe labelling and hybridisation to Southern blot

The DNA transferred onto a nylon membrane was visualised by the use of a radioactively labelled probe. The method used to label the probes consisted of the use of random hexamers and the inclusion of [α-^{32}P] dCTP by the DNA polymerase Klenow fragment (Feinberg and Vogelstein, 1984). The amount of probe labelled varied from 15 to 20 ng and 5 ng for marker DNA. The labelling reaction was incubated overnight at room temperature and stopped by the addition of 1.6 volumes of 20 mM NaCl, 20 mM Tris-HCl pH 7.5, 2 mM EDTA and 0.25% SDS solution. The probe was ethanol-precipitated using 90 µg of high molecular weight salmon sperm DNA as a carrier. Probes were dissolved in 0.5 ml water and boiled for 5 min to denature the DNA before use.

Southern blots were pre-hybridised for at least 30 min at 65°C (or 60°C for telomere specific probes) in a phosphate-SDS solution (7 % SDS, 0.5 M Na_2HPO_4 pH 7.2, 1 mM EDTA). The pre-hybridisation solution was replaced with fresh solution and the probe was added at this stage. Hybridisation was carried out overnight in a hybridasation oven (Hybaid, Thermo scientific). Blots were washed 2 to 3X 20 min under low stringency conditions (1X SSC, 0.1% SDS). The blots were kept damp and were exposed to autoradiograph films (Fuji RX100 X-ray film) at -80°C with an intensifying screen or alternatively to storage phosphor screens and visualised using a phosphorImager device.

2.2.11 Western blotting

Western blotting was used to identify a specific protein from total protein cell extracts. The total protein extract was separated by SDS-polyacrylamide gel electrophoresis and the proteins were transferred to a nitrocellulose membrane where they were probed using antibodies.

2.2.11.1 SDS-polyacrylamide gel electrophoresis of proteins

Electrophoresis of proteins was carried out in SDS-polyacrylamide gels to ensure dissociation of proteins and thus facilitating their migration on gel. In order to increase the resolution of the gel, the electrophoresis was carried out with a discontinuous buffer system. The samples migrate firstly through a 3% stacking acrylamide gel which was prepared by adding 250 µl of 30% acrylamide/0.8% bis-acrylamide (National Diagnostics), 500 µl of

stacking buffer (0.5 M Tris base, 0.4% SDS, pH was adjusted to 6.8 with HCl) and made up to 2 mls with water. Subsequently, the samples were separated using a 7.5% acrylamide separating acrylamide gel prepared as follows: 1.5 mls 30% acrylamide/0.8% bis-acrylamide, 1.5 mls separating buffer (1.5 M Tris base, 0.4% SDS, pH was adjusted to 8.8 with HCl) and made up to 6 mls with water. The gels were polymerised by the addition of 75 µl of 10% ammonium persulfate and 5 µl of TEMED (7.5% acrylamide gels separate proteins from ~20 to 150 kDa).

2.2.11.2 Sample preparation and gel electrophoresis

Leammli loading buffer (for 5 X: 50% glycerol, 0.025% bromophenol blue, 0.05% β-mercaptoethanol, 0.1% SDS, 0.3 M tris pH 6.8) was added to the samples and boiled for 3 to 5 min. Approximately 30 to 35 µg of proteins were loaded in each well and the gel was run in an electrophoresis chamber (Mini-Protean 3 Cell, BioRad) at 150 V for 1 hr (until dye reached the bottom of the gel). The buffer used to run the gel was kept as a 10X stock made of 220 mM Tris, 1.92 glycine, 1% SDS. In every gel, a molecular size marker was loaded (ECL DualVue Western Blotting Marker, GE Healthcare).

2.2.11.3 Transfer of proteins onto a membrane

Protein transfer was carried out in semi-dry conditions onto a nitrocellulose membrane (Hibond-ECL BioRad). The membrane and the gel were incubated for 10 min in transfer buffer (25 mM Tris, 192 mM glycine, 20% ethanol) and the transfer was made on a semi-dry transfer unit (BioRad) at 1 mAmp per cm^2 for 1 hr. The membrane was washed in water and kept damped at 4°C until antibody staining.

2.2.11.4 Protein detection by antibody staining

In this work, the detection of proteins was carried out by antibody coupled to horseradish peroxidase using the ECL (enhance chemiluminescence) technique. ECL uses the horseradish peroxidise enzyme to oxide (by hydrogen peroxide) a chemiluminescence substrate such as luminol to produce visible light.

The Western blot membrane was blocked with 5% milk powder dissolved in PBS/0.1% Tween 20 for at least 30 min. The membrane was then incubated with the primary antibody in 5% milk powder, PBS/0.1% Tween 20 for at least 1 hr rotating in a tube. MSH2 mouse antibody (Calbiochem, Germany) was used at 1:200 dilution while β-Actin mouse antibody (Abgene, UK) at 1:12,000. After antibody incubation, the membrane was washed 4

times with 0.1% PBS/Tween20 over a period of 30 min. The secondary antibody was an anti-mouse horseradish peroxidise antibody (raised in sheep; GE healthcare, UK), it was incubated with the Western blot for 45 min at a dilution of 1:20,000. The membrane was washed as before and it was developed using the ECL plus western blotting detection system (GE healthcare, UK). The signal was visualised by exposing the membrane to a film from 30 sec to 5 min.

2.2.12 DNA sequencing

The method used for sequencing was the dideoxy-mediated chain terminator method that consists in the use of dideoxynucleside triphosphates (which lack a 3'-hydroxyl residue) to terminate the DNA synthesis in a random base-specific way.

The fragment of DNA to be sequenced was PCR amplified and gel purified using the MinElute Gel Extraction Kit (Qiagen). DNA sequencing was carried out using an Applied Biosystems 3730 DNA sequencing system with 1 µl of the Big Dye™ Terminator Cycle Sequencing Ready Reaction Kit (version 3.1), 1.5 µl of 5x Big Dye™ Terminator buffer, 20 – 30 ng of template PCR product per kb of sequence and 3.2 pmol of sequencing primer in a final volume of 10 µl. The reaction was cycled 25 times as follows; 10 sec at 96°C, 5 sec at 50°C and 4 min at 60°C. After the sequencing reaction, the sample was purified by adding an equal volume of dH_2O, 2 µl of 2.2% SDS and incubated for 5 min at 98°C and 10 min at 25°C using a PTC200 DNA Engine thermal cycler (MJ Research, USA). Excess of dye was removed by placing the reaction onto spin columns (Performa DTR gel filtration cartridges, Edge Biosystems). Gel running was carried out in the Protein and Nucleic Acid Laboratory (PNACL), University of Leicester. The sequence was analysed using the Chromas software (version 1.45).

2.2.13 Restriction Fragment Length Polymorphism (RFLP) Analysis

2.2.13.1 Analysis of the Xp/Yp telomere-adjacent sequence

In order to carry out analysis of specific telomeres, the haplotype of the Xp/Yp telomere-adjacent sequence was determined using RFLP analysis on a 500bp telomere adjacent fragment generated with TSK8G and TSK8C primers. This amplicon contains seven SNPs which are almost in complete linkage disequilibrium, generating three different haplotypes: A, B and C (Baird *et al*, 1996). For the purpose of this work, two SNPs were assayed (-415 and -427), which are enough to determine the haplotype in the DNA sequence immediately adjacent to the Xp/Yp telomere. The –415 SNP is polymorphic for C or T;

when a C base is present at this position, two *Mbo*II fragments are generated (314 bp + 184 bp) but the presence of a T at the –415 base creates another cutting site in the 314 bp fragment, producing a 268 bp and a 46 bp fragments (Baird *et al*, 1996). The –427 SNP is polymorphic for a G or A. The presence of a G base at this position creates a *Taq*I restriction site, which cuts the 500 bp amplicon into two: a 448 bp and a 50 bp segment. The products were resolved by electrophoresis on 1.5% agarose gels.

2.2.13.2 Analysis of the 12q telomere-adjacent sequence

Four polymorphic sites were assayed in order to determine the haplotype of the sequence adjacent to the 12q telomeres. These sites are: -1036 (*Taq*I +/-), -1812 (*Alu*I +/-), -473 (*Bsm*I +/-), -554 (*Kpn*I +/-) and a (659-630) insertion/deletion (the bases are numbered with respect to their distance from the start of the telomere). The 12q telomere-adjacent region was PCR amplified with primers pKSRV2D and 12qAReverse to assay the polymorphic sites -1036 (*Taq*I +/-) and -1812 (*Alu*I +/-). This amplicon was generated with the following conditions: 96°C for 20 sec, 63°C for 30 sec and 70°C for 1 min for 33 cycles. The PCR products were digested with the enzymes *Taq*I and *Alu*I and the products were resolved by electrophoresis on 4.5% and 1.5% agarose gel respectively.

Polymorphic sites -473 (*Bsm*I +/-), -554 (*Kpn*I +/-) and the -(659-630) insertion/deletion were assayed from amplicons generated with 12qB and 12qA primers (Baird *et al.*, 2000). The amplicons were digested with the appropriate enzyme and the products were resolved in agarose gels from 1.5% to 3.5% according to the size of the fragments generated.

2.2.14 Detection of the polymorphic 16p/16q telomere

To determine the presence or absence of the polymorphic telomere on chromosome 16 (Coleman *et al.,* 1999), genomic DNA was amplified with the primers Nitu14eA and TelC (this primer anneals to the G-rich strand of the telomere) under the following conditions: 96°C for 20 sec, 68°C for 30 sec and 70°C for 5 min for 20 cycles. The amplified products were resolved on 1.2% agarose gels. The DNA products were blotted onto a nylon membrane and hybridised in phosphate-SDS solution at 65°C overnight to a radioactively labelled probe (Coleman *et al.*, 1999). The probe (10 ng) was generated by PCR of genomic DNA from samples containing the 16p/q telomere using primers Nitu14eA together with Nitu14eB and labelled with 2 μCi [α-^{32}P] dCTP.

Chapter 3

Telomere Mutation Analysis in Mismatch Repair Deficient Cell Lines

3.1 Introduction

3.1.1 Stability of microsatellites is controlled by mismatch repair

The first evidence that suggested that DNA mismatch repair is involved in the stability of microsatellites came from the observation that a proportion of sporadic colorectal cancers and cancers in HNPCC patients harbour instability at short-tandem repeats (Ionov *et al.,* 1993; Aaltonen *et al.*, 1993). Since then, different studies have shown that germ-line mutations in the *MSH2, MSH6, MLH1* and *PMS2* genes are associated with HNPCC and thus linked to microsatellite instability (MSI) (Bronner *et al.*, 1994; Papadopoulos *et al.*, 1994; Miyaki *et al.*, 1997; Hendriks *et al.*, 2006).

Studies in yeast have shown that mutations in *MSH2*, *MLH1* or *PMS1* lead to 100- to 700-fold increase in the mutation rate of microsatellites most probably produced by replication slippage events (Strand *et al.*, 1993). Similar studies have been performed in colorectal carcinomas and it has been shown that MSI is present in coding and non-coding microsatellites, however there is a high mutation frequency in tracts within growth related genes such as *BAX* and *TGFβRII* with about 40% and 90% mutation frequency respectively (Woerner *et al.*, 2005). It is proposed that high mutation frequency in such genes reflects selective pressure as mutations in these genes confer growth advantage to those cells. However, MSI has also been found in non coding sequences, for example about 51% of colon carcinomas showed MSI in a tract of 9 thimidines found within an intron of the *SEMG1* gene (Hienonen *et al.*, 2005) and a mutation frequency of about 80% has been found in a microsatellite found in an intron of the *ATM* gene (Ejima *et al.*, 2000). Further evidence in colorectal cancer cell lines showed that MSI occurs in coding and non-coding sequences with a mutation rate as high as 10^{-2} in a cell line with mutations in the *MLH1* gene (HCT116) compared to a normal cell line that showed a mutation rate of 10^{-8} per cell per generation (Bhattacharyya *et al.*, 1994).

Mutation rates have also been study at the hypoxanthine-guanine phosphoribosyl transferase (*HPRT*) gene. This gene is 44 kb in length and is composed of nine exons. In addition, the *HPRT* gene contains a run of six guanines in exon three. Bhattacharyya *et al* (1995) studied the mutation rate at this gene in three different MMR deficient cell lines, HCT116 that is defective in the *MLH1* gene and the *MSH6* defective cell lines, DLD-1 and HCT15. The mutation rate at the *HPRT* gene was 10^{-5} per cell per generation in the three MMR defective cell lines, however there was at least 1000-fold increase in the mutation rate compared to a normal cell line (MRC5). Although some of the mutations were found in the run of six guanines, other insertions, deletions and base substitution were found across the whole gene. These results indicate that defects in MMR genes not only affect microsatellites but also have an overall effect in spontaneous mutation rates.

3.1.2 Chromosomal instability in colon carcinomas

As mention above, a proportion of sporadic colon carcinomas (around 10%) show MSI that is related to defects in the MMR system. The majority of colon carcinomas, on the other hand, are microsatellite stable (MSS) but such carcinomas contain high levels of chromosomal aberrations such as duplications, deletions and translocations of chromosomes. Studies in cell lines have showed that the rate of chromosome gains or losses is 10^{-2} per population doubling per chromosome for MSS lines, while MSI cell lines presented very few alterations, with the same frequency of gains and losses as a lymphoblast control cell line. Abdel-Rahman *et al* (2001) investigated the chromosomal rearrangements in 17 cell lines (MSS and MSI) derived from colorectal carcinomas. They found that all of the MSS cell lines had abnormalities in the chromosome number (gains and losses), numerous chromosome rearrangements (mainly unbalanced translocations) and multiple trisomies. Six out of eight MSI cell lines (such as LoVo, DLD-1, Vaco5 and HCT116 which were used in this work) presented few abnormalities in chromosome number or breakpoints. Two other cell lines with a MSI phenotype showed a different pattern of instability. The LS411 cell line, for example, displayed characteristics of MSS cells, with a near-triploid karyotype and multiple translocations. Another cell line (HCA7) had a near-diploid chromosome number as MSI but it showed multiple breakpoints and translocations. This shows that MSI and chromosomal instability can coexist in the same cell line.

Chromosome instability has been associated with mutations in oncogenes such as the *TP53* gene. It has been reported that the lack of p53 in mouse embryonic fibroblast cells causes the production of multiple copies of centrosomes, leading to unequal chromosome

segregations and then to chromosome instability (Fukasawa *et al.*, 1996). However, most of the colorectal carcinomas with MSI have wild type copies of the *TP53* gene. Thus, it is thought that the p53 protein is important in maintaining chromosomal stability in MSI tumours. Studies in cell lines showed that among fourteen colorectal cancer cell lines (five MSS and nine MSI) all the MSS cell lines had mutations at the *TP53* gene and show chromosome instability. Among the nine MSI cell lines, five did not show *TP53* mutations but the four that did, had stable karyotypes. Therefore is also possible that in some MSI cell lines, chromosome stability could be maintained by mechanisms independent of p53 (Eshleman *et al* 1998; Mancuso *et al.*, 1997).

3.2 Aims

MSI is known to be involved in the development of HNPCC and some sporadic colorectal cancers, and defects in the MMR pathway have a major role in this process. It has been shown that MMR defects in tumours and cell lines induce microsatellite instability. Previously, it was seen that telomeres mutate *in vivo* in sporadic colorectal tumours and some of those carcinomas were known to contain mutations in the *MSH2* gene (Pickett *et al.*, 2004). Therefore instability at telomeres in colon carcinomas might be linked to defects in the MMR pathway.

In this work, colorectal cancer cell lines with defects in the *MSH2*, *MSH6* or *MLH1* genes were assayed to determine the contribution of each gene to telomere instability. Furthermore, to detect mutations that may occur at low frequency within a population of cells (*i.e.* using bulk DNA) a cell cloning approach was used. Around one hundred clones were analysed by TVR-PCR in order to identify the potential effects that different MMR genes had on telomeres.

3.3 Results

3.3.1 Cell lines

The cell lines used in this work were provided by Ian Tomlinson (Cancer Research UK, London). Table 3.1 summarises the MSI status and some of the mutations in mismatch repair, oncogenes and tumour suppressor genes found in these cell lines (Eshleman *et al.*, 1998; Gayet *et al.*, 2001; Abdel-Rahman *et al.*, 2001). The microsatellite instability status of

most cell lines was determined by analysing the stability of mono and di-nucleotide repeats (Cottu *et al.*, 1996; Hoang *et al.*, 1997; Efstathiou *et al.*, 1999). All the cell lines (with the exception of LS411) have a stable karyotype (almost diploid) which is characteristic of most MSI cell lines and tumours (Abdel-Rahman *et al.*, 2001).

3.3.2 Haplotype analysis at the Xp/Yp and 12q telomere-adjacent region

In order to analyse telomere mutation frequency using TVR-PCR, it is advantageous to analyse telomeres that are heterozygous at the subtelomeric region. Amplification of telomeres by TVR-PCR in heterozygous samples generates cleaner telomere maps facilitating the detection of mutations. Therefore, haplotype analysis of Xp/Yp and 12q telomeres was carried out in the cell lines showed in table 3.1. Briefly, for the Xp/Yp telomere, restriction fragment length polymorphism (RFLP) analysis was carried out on a 500 bp fragment generated with TSK8G and TSK8C primers. This amplicon contains seven SNPs which are almost in complete linkage disequilibrium, generating three different haplotypes: A, B and C (Baird *et al*, 1995). The amplicon generated was digested with *Mbo*II and *Taq*I restriction enzymes as indicated in the Materials and Methods section 2.2.13.1. Among the seven cell lines analysed, five were heterozygous for haplotypes in the Xp/Yp telomere-adjacent region (Table 3.2). These cell lines are potentially useful for subsequent telomere analysis.

Analysis at the 12q telomere adjacent region was also carried out to determine heterozygosity at this region. The 12q subtelomeric region analysed (1870 bp) contains twenty polymorphic sites and a 30 bp insertion/deletion and similar to the Xp/Yp telomere adjacent region, these SNPs are almost in complete linkage disequilibrium creating three common haplotypes: A, B and A1 (Baird *et al.*, 2000). Three polymorphic sites that alter the digestion pattern of restriction enzymes were tested in order to determine the haplotype in the region adjacent to the 12q telomere (Table 3.3). Two of the cell lines (LS411 and HCT15) were heterozygous for different haplotypes in the 12q telomere adjacent sequence but the other samples were homozygous thus no further analysis was carried out in these samples.

Cell line	SEX/AGE	MSI	MMR	Mutations			Karyotype
				K-RAS	*TP53*	*APC*	Chromosome number
HCT15	M?	+	*MSH6*	Codon 12	Codon 241	Yes	46
LS411	M33	+	*MLH1*	None	Codon 126	Yes	70-76
LS180	F58	+	*MLH1*	Codon 12	None	None	45
VACO5	F78	+	*MLH1*	None	Codon 282	Yes	47
LoVo	M56	+	*MSH2*	Codon 13	None	Yes	47-50
HCT116	M?	+	*MLH1*	Codon 13	None	None	46
HCT116 +3 *	M?	-	RESTORED *MLH1*	Codon 13	None	None	46

Table 3.1 Mutations of oncogenes, tumour suppressor and MMR genes identify in colorectal cancer cell lines

All the cell lines described in this table were derived from colorectal carcinomas. Apart from LS411 cell line, all the lines have stable karyotypes with few translocations events. * HCT116 +3 cell line was generated by transfection of chromosome 3, which contains a wild-type copy of the *MLH1* gene, into HCT116 cell line. Most of the information shown in this table can be accessed through the link http://www.cephb.fr/gaccc.

Cell line	Genotype		Haplotype
	-415 *Mbo*II	-427 *Taq*I	
HCT15	C/T	A/G	A/B
LS411	C/C	G/G	A/A
LS180	T/T	A/A	B/B
VACO5	C/T	G/G	A/C
LoVo	C/T	A/G	A/B
HCT116	C/T	G/A	A/B
HCT116 +3	C/T	G/A	A/B

Table 3.2 Haplotype analysis at the Xp/Yp telomere using *Mbo*II and *Taq*I sites to assess the SNPs at positions -415 and -427 respectively. The SNPs are numbered relative to the start of the telomere such that -1 is the first base internal to the telomere.

Cell line	Genotype				Haplotype
	-1036 *Taq*I	(30bp in/del)	-554 *Kpn*I	-437 *Bsm*I	
HCT15	C	+	A	A	A/Δ allele
LS411	C/T	+/-	A/A	A/G	A/B
LS180	C/C	+/+	A/A	A/A	A/A
VACO5	T/T	-/-	A/A	G/G	B/B
LoVo	C/C	-/-	A/A	A/A	C/C
HCT116	C/C	+/+	A/A	A/A	A/A
HCT116 +3	C/C	+/+	A/A	A/A	A/A

Table 3.3 Haplotype analysis in the 12q telomere-adjacent DNA sequence. The SNPs positions and the enzymes used to detect them are shown. The + or – at the 30 bp insertion/deletion polymorphism indicates presence or absence (respectively) of the sequence. The haplotypes were determined according to Baird *et al.*, 2000.

3.3.3 Detection of the polymorphic 16p/16q telomere and another telomere known as CB54

Cell lines were analyzed for the presence of a polymorphic telomere on chromosome 16 which is found in approximately 6% of the 16q and 2% of the 16p arms. PCR was carried out with primers Nitul4eA that anneals to the telomere-adjacent DNA of this polymorphic sequence together with primer TelC that anneals to the G-rich strands of the TTAGGG

repeats (further details can be found in section 2.2.14). The PCR products were detected following southern blot hybridisation to a probe made from the telomere-adjacent sequence of these telomeres (Coleman *et al.*, 1999). None of the samples analysed revealed the presence of the polymorphic 16p/q telomere (Figure 3.1) by southern blotting, thus no further analysis was done on that telomere.

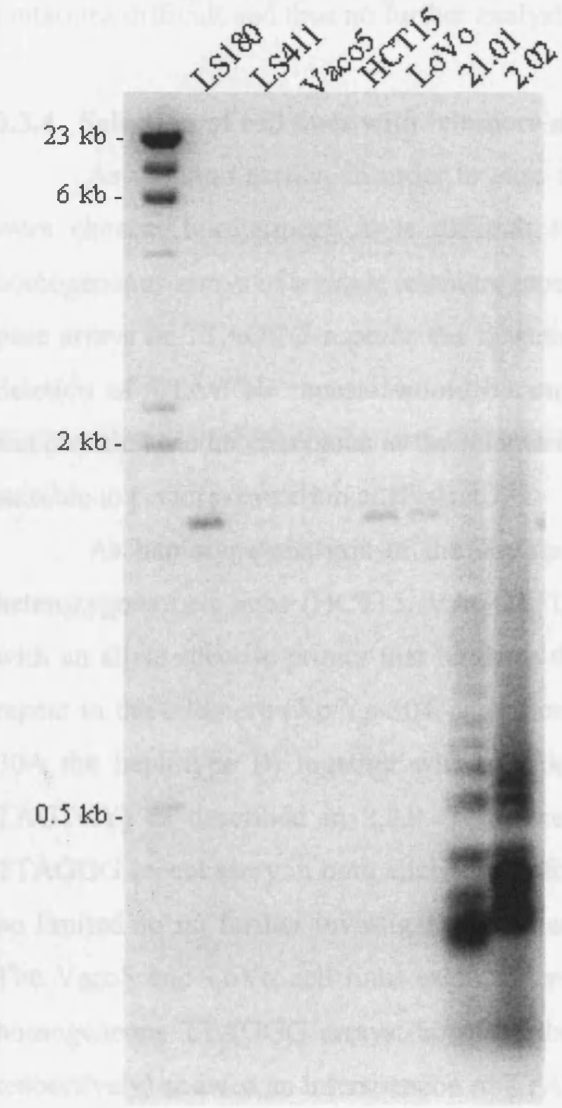


Figure 3.1 Southern blot showing the presence or absence of the polymorphic telomere on chromosome 16p/q. The two CEPH DNA samples (21.01 and 2.02) were used as controls as they each contain 1 copy of this telomere. The shortest product expected is 200 bp.

Another polymorphic telomere was isolated from a cell line (CB00054) by Helen Varley. This telomere-adjacent DNA is found next to a telomere in approximately 11% of the Caucasian population and it shows 93% sequence homology to a region located 26 kb from the start of the 7q telomere. To date, the chromosomal location of this polymorphic telomere has not been determined so it was referred to as the CB54 telomere.

To determine whether any of the cell lines contained the CB54 telomere, a PCR was carried out using primer CB38A-F together with Tel C. The PCR product was resolved in an agarose gel, blotted and hybridised to a probe made from the CB54 telomere adjacent sequence amplified from the CB00054 cell line. Initial screening showed that LoVo cell line contains one copy of the CB54 telomere (Figure 3.2). Further analysis was carried out by performing TVR-PCR in LoVo cell line; however the results showed that the CB54 telomere in LoVo is composed of a homogeneous array of TTAGGG repeats, making the detection of mutations difficult and thus no further analysis was carried out in this telomere end.

3.3.4 Selection of cell lines with telomere maps suitable for mutation detection

As mention earlier, in order to ease the mutation detection, heterozygous cell lines were chosen. Furthermore, it is difficult to detect mutations in cell lines that contain homogeneous arrays of a single telomere repeat. In other words, if a telomere is composed of pure arrays of TTAGGG repeats, the detection of simple mutation such as the insertion or deletion of TTAGGG repeats would be impossible. Thus, TVR-PCR was carried out in samples that are heterozygous at the telomere-adjacent DNA in order to find telomere maps suitable to perform mutation analysis.

As haplotype analysis of the Xp/Yp telomere-adjacent region had identified three heterozygous cell lines (HCT15, VACO5, LoVo) TVR was carried out on these samples with an allele specific primer that binds to the polymorphism that lies 30 bp from the first repeat in the telomere (Xp/Yp-30T amplifies the haplotype A-associated allele and Xp/Yp-30A the haplotype B) together with a telomere mapping primer (TAGTelW, TAGTelX, TAGTelY) as described in 2.2.9. Telomere maps from HCT15 showed a homogeneous TTAGGG repeat array in both alleles, therefore mutation detection in these telomeres would be limited so no further investigation was carried out on the Xp/Yp telomeres in HCT15. The Vaco5 and LoVo cell lines each presented one haplotype A-associated telomere with homogeneous TTAGGG arrays; however the other alleles (haplotype C and haplotype B, respectively) showed an interspersion of TTAGGG with sequence variant repeats.

In a similar way, telomere maps were generated at 12q for the heterozygous cell lines (LS411 and HCT5). TVR-PCR at 12q was conducted with allele specific primers (12q-197A, which amplifies from haplotype A, 12q-197G which amplifies from haplotype B and 12qnull3 for the 12q-deletion allele) together with a telomere mapping primer. The telomere and variant repeat distribution patterns for both alleles of both cell lines were suitable for analysis.

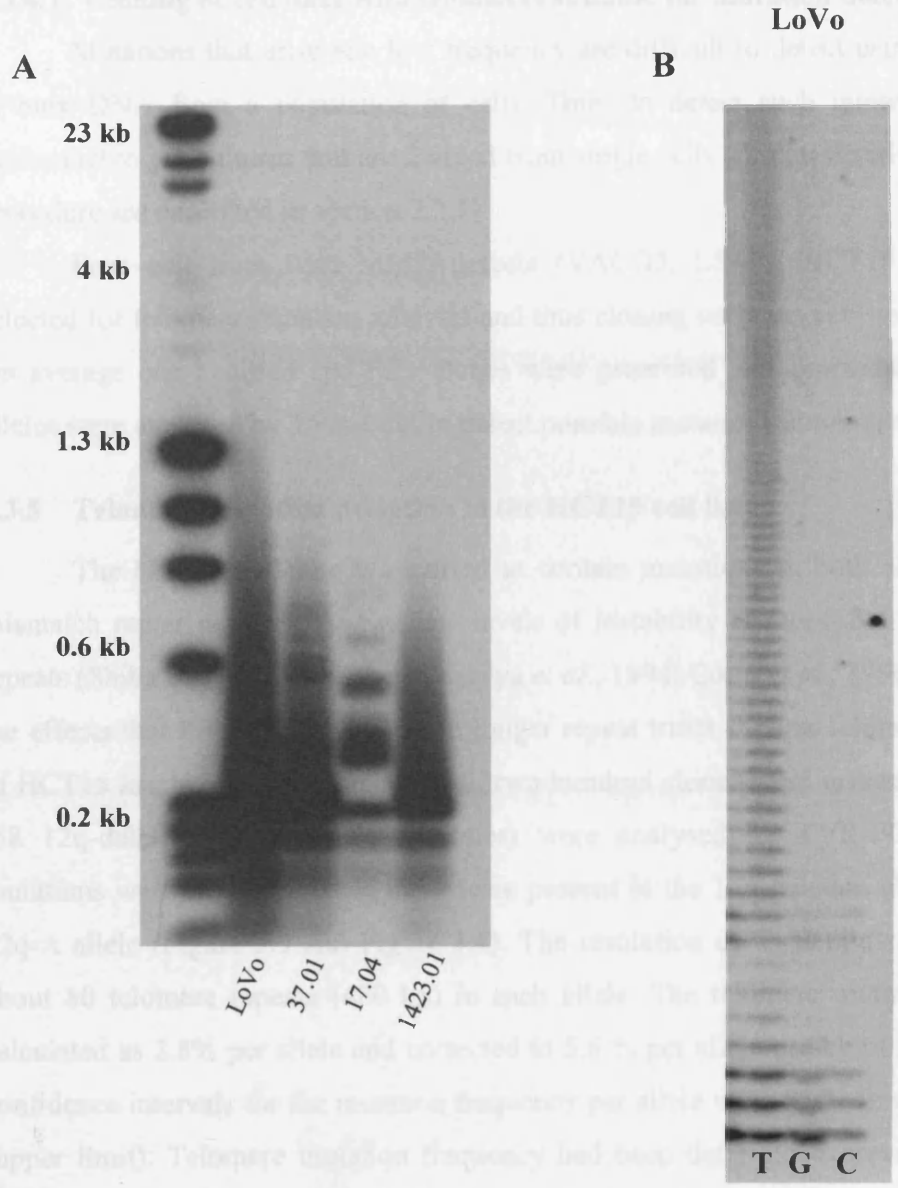


Figure 3.2 Southern blot detection of PCR products from the CB54 telomere. Telomere amplicons detected in the LoVo cell line showed the presence of the CB54 telomere (*A*). The CEPH 37.01, 17.04 and 1423.01 DNA samples were used as positive controls. (*B*) TVR-PCR of LoVo DNA using a primer that anneals to the CB54 telomere-adjacent sequence together with a telomere mapping primer that anneals to a specific variant telomere repeat was used. Telomere mapping primers that anneal to the TTAGGG (T), TGAGGG (G) and TCAGGG (C) repeats were used.

3.3.4.1 Cloning of cell lines with telomeres suitable for mutation detection

Mutations that arise at a low frequency are difficult to detect using PCR techniques in bulk DNA from a population of cells. Thus, to detect such mutations, cloning was performed to get cultures that are derived from single cells (further details about the cloning procedure are described in section 2.2.1).

Four cell lines with MMR defects (VACO5, LS411, HCT15 and LoVo) were selected for telomere mutation analysis and thus cloning on those cell lines was carried out. On average one hundred and fifty clones were generated and approximately one hundred alleles were analysed by TVR-PCR to detect possible mutations at telomeres.

3.3.5 Telomere mutation detection in the HCT15 cell line

The HCT15 cell line is reported to contain mutations in both alleles of the *MSH6* mismatch repair gene and shows low levels of instability at mono but not at dinucleotide repeats (Shibata *el al.*, 1994; Bhattacharyya *et al.*, 1994; Cottu *et al.*, 1996). In order to study the effects that this gene may have on longer repeat tracts such as telomeres, TVR analysis of HCT15 line was carried out. Overall, two hundred clones were generated and 143 alleles (68 12q-deletion and 75 12q-A alleles) were analysed by TVR-PCR. Four telomere mutations were found, three of them were present in the 12q-deletion allele and one in the 12q-A allele (Figure 3.3 and Figure 3.4). The resolution of TVR allowed the screening of about 80 telomere repeats (480 bp) in each allele. The telomere mutation frequency was calculated as 2.8% per allele and corrected to 5.6 % per allele per kb of sequence. The 95% confidence intervals for the mutation frequency per allele were 0.9% (lower limit) and 7.2% (upper limit). Telomere mutation frequency had been determined previously by the same cloning approach in primary and telomerase positive cell lines and the mutation frequency was 0 in both cell types (0/504 mutants; confidence limits: 0 - 0.64% and 0/344 mutants; confidence limits: 0 – 0.94% respectively; Varley *et al.*, 2002). The telomere mutation frequency found in this cell line was significantly different from primary cell lines (Fisher's exact test, two-tailed, $P = 0.0023$) and from telomerase positive (MMR+) cell lines ($P = 0.0072$, Table 3.4).

The mutations observed are likely to be originated through an intra-allelic mechanism such as replication slippage, unequal telomere sister chromatid exchanges or gene conversion events. One mutation has resulted in an insertion of five TGAGGG type repeats (clone 33) and the remaining three were repeat-type changes (clones 11, 61 and 67; Figure 3.3 and Figure 3.4).

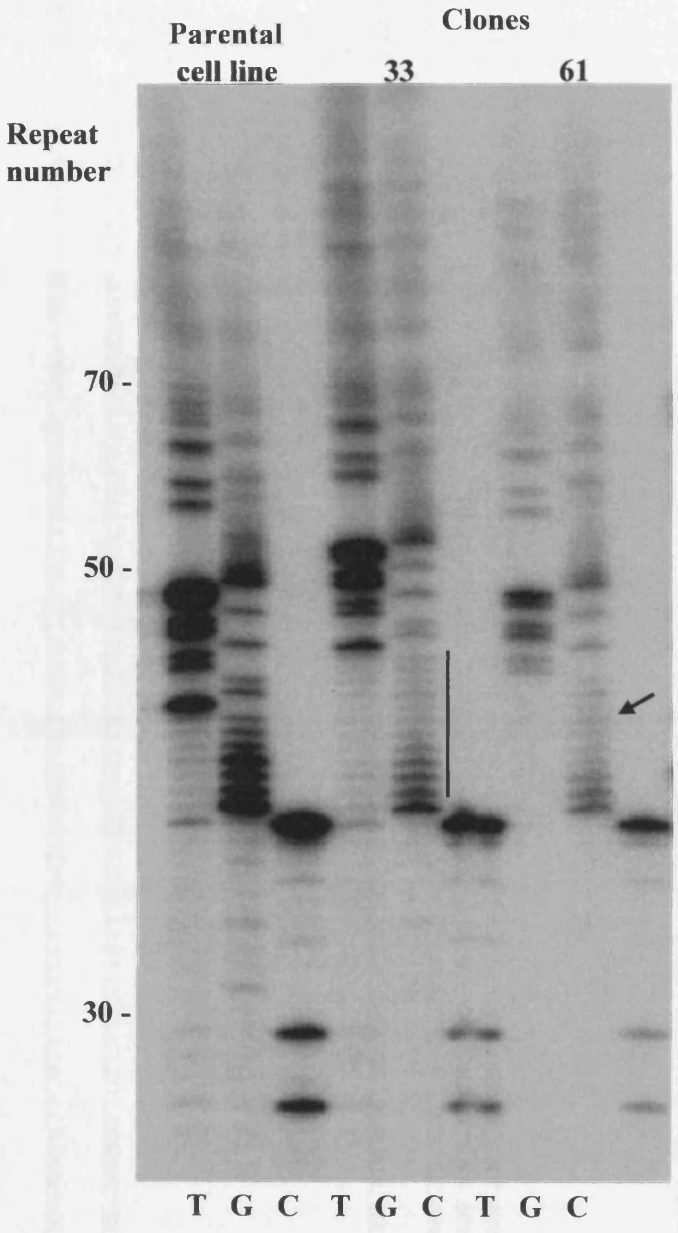


Figure 3.3 Telomere maps at the 12q deletion allele in the HCT15 cell line and clones. TVR-PCR at the 12q deletion-associated telomere in HCT15 progenitor cell line and from two clonal populations are shown. Clone 33 shows an insertion of five G-type telomere repeats (bar) and clone 66 shows change of repeat type from a G-type to a T-type repeat (arrow). The first ~30 repeats are not shown. Repeat types: T (TTAGGG), G (TGAGGG), C (TCAGGG).

HCT15 telomere maps

A) Haplotype A at telomere 12q

B) 12q deletion allele

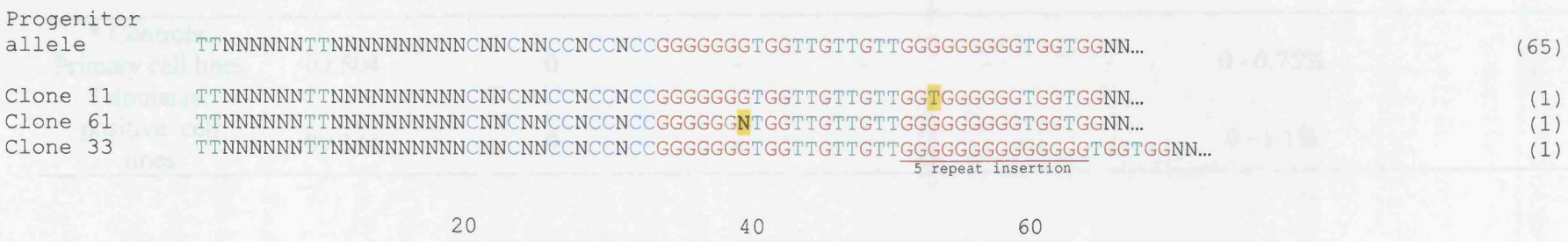


Figure 3.4 Comparison of the 12q associated telomere mutants from HCT15 clones

The map of each progenitor allele was determined by TVR-PCR analysis of bulk DNA from a non-clonal HCT15 population. The progenitor telomere map was compared with the maps derived from clones that showed mutations. Three of the mutations are repeat type changes (highlighted in yellow) and one is an insertion of five TGAGGG telomere repeat-types (underlined). In parenthesis at the end of each telomere map is shown the number of times the progenitor and the mutant alleles were observed. T, TTAGGG; G, TGAGGG; C, TCAGGG; N, repeats that do not amplify with the primers used.

Cell line (MMR gene defect)	Mutants / Alleles analyzed	Mutation frequency (per kb)	Type of Mutations				95% confidence intervals	Fisher's Exact test ‡ (Primary; telomerase positive cell line)
			Insertions	Deletions	Base changes	Unknown	Low - High	
HCT15 (*MSH6*)	4 / 143	2.8% (5.6%)	1		3		0.9% - 7.2%	$P = 0.0023$; 0.0072
LS411 (*MLH1*)	2 / 104	1.9% (3.9%)		1		1	0.1% - 6.8%	$P = 0.03$; 0.05
Vaco5 (*MLH1*)	3 / 110	2.7% (6.4%)	1	2			0.6% - 8.1%	$P = 0.006$; 0.014
LoVo (*MSH2*)	7 / 93	7.5% (16%)	1	5		1	3.5% - 15%	$P < 0.001$
Total	16 / 450	3.6% (7.7%)	3	8	3	2	2.2% - 5.7%	$P < 0.001$; 0.0015
* Controls: Primary cell lines	0 / 504	0	-	-	-	-	0 - 0.75%	
Telomerase positive cell lines	0 / 344	0	-	-	-	-	0 - 1.1%	

Table 3.4 Mutation frequencies of MMR cell lines

The mutation frequency per allele was corrected for kb of sequence and given as a percentage. Two mutations could not be classified as insertion, deletions or base change repeats thus they were called unknowns. The 95% confidence intervals were estimated using the mutation frequency per allele. ‡Telomere mutation frequency of MMR defective cell lines was compared to the mutation frequency of primary cell lines and telomerase cell lines. The *P* values were obtained from the Fisher's exact test (two-tailed). * Control data was obtained from Varley *et al.*, 2002.

3.3.6 Telomere mutation detection in the LS411 cell line

LS411 has a missense mutation in codon 226 of one allele of the *MLH1* gene. This mutation was caused by a single base-pair change of CGG for TGG (Arg for Trp) while the promoter of the remaining copy of *MLH1* is methylated (Wheeler *et al.*, 1999). The same authors confirmed by western blotting the lack of Mlh1 protein in this cell line.

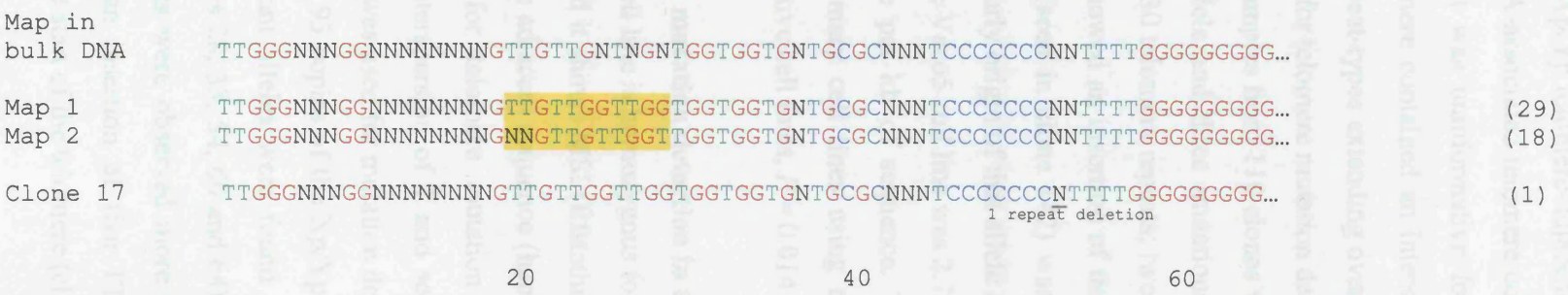
LS411 cell line was cloned and 102 alleles (48 12q-A and 56 12q-B alleles) at the 12q telomere were screened for mutations. In bulk DNA from the LS411 cell line the haplotype A-associated telomere showed a telomere map that suggested the presence of two different populations of cells that have different telomere maps. This observation was confirmed when analysing the A-haplotype alleles in clones. It was possible to identify two different types of maps, with only a small region (11 repeats) being different between them (Figure 3.5). Map 1 was seen in approximately 62% of the alleles while map 2 was seen in the remaining 38%. Therefore the cell line is mosaic for two different haplotype A-associated telomere maps, presumably because a telomere mutation arose earlier in the establishment of this cell line. This eleven repeat difference between the two maps was difficult to score as an insertion, deletion or repeat-type change mutation. One further mutation was observed in clone 17, which shows a deletion of one N-type repeat in the telomere associated with the 12q-A haplotype (Figure 3.5).

In addition to the mutation found in the 12q-A haplotype, six repeat type variations (appearance or disappearance of telomere repeats) were found in the 12q-B haplotype. These variations lie in a region where the interspersion of TTAGGG and variant repeats is unclear. This cell line has been assessed for chromosome stability and it contained about 75 chromosomes including at least four copies of the 12q chromosome arm (Abdel-Rahman *et al.*, 2001). Thus, the uncertainty in this region of the 12qB map may arise from amplification of multiple alleles. Another possibility is that the changes arose from telomere-repeat amplification from elsewhere in the genome. The 12q and 7q telomere adjacent DNA sequence share 97.9% sequence similarity, thus it is possible that the extra-bands observed came from the 7q telomere. Therefore, it was not possible to conclude whether these possible repeat-type changes were real mutations.

Hence, the estimate of the telomere mutation frequency in the LS411 cell line (omitting the unverified changes in the 12qB associated telomere) was 1.9% per allele (2/104; 95% CI, 0.1% - 7.2%) and corrected to 3.9% per kb per allele. The telomere mutation frequency found in this cell line was significantly different from primary cell lines using Fisher's exact test, two-tailed, $P = 0.03$ (Table 3.4).

LS411 telomere maps

12q haplotype A-associated telomere



12q haplotype B-associated telomere

Figure 3.5 Comparison of the 12q associated telomere mutants from LS411 clones

Telomere maps associated with 12q haplotypes A and B are shown. The cloned DNA samples contained either 12q map 1 or map 2 showing the cell line is mosaic for these two maps that differ in the 11-repeat region highlighted in yellow. Clone 17 contained map 1 with an additional single repeat deletion mutation. The position of the deletion in clone 17 is underlined. In parenthesis is shown the number of times the progenitor and the mutant alleles were observed. No mutations were found at the 12q haplotype-B associated telomere. T, TTAGGG; G, TGAGGG; C, TCAGGG; N, repeats that do not amplify with the primers used.

3.3.7 Telomere mutation detection in the Vaco5 cell line

Vaco5 cell line contains a mutation in codon 409 in one allele and promoter methylation of the second allele of the *MLH1* gene (Wheeler *et al.*, 1999; Bhattacharyya *et al.*, 1994). The Xp/Yp telomere-adjacent region is heterozygous for the A and C haplotypes. The haplotype A-associated telomere comprised a homogeneous array of TTAGGG repeats and therefore it was uninformative for mutation detection. However, the haplotype C-associated telomere contained an interspersed pattern mainly of the T (TTAGGG) and G (TGAGGG) repeat-types extending over the proximal 80 repeats (480 bp) and therefore this allele was used for telomere mutation detection.

DNA samples from 110 clones were analyzed by TVR-PCR at the Xp/Yp haplotype C-associated allele and three mutations were detected. All of the mutations were found within the first 30 telomere repeats; two were deletions of one T-type repeat while the other mutant allele showed an insertion of three G-type repeats (Figure 3.6). One of the deletion mutant alleles (seen in clone 142) was identified in more than one clonal DNA sample, suggesting an early origin of that allele during the propagation of the cell line. The mutation frequency in the Vaco5 cell line was 2.7% per allele (95% CI; 0.6% - 8.1%) and corrected to 6.4% per allele per kb of sequence. The telomere mutation frequency was statistically different to primary cell lines using the Fisher's exact test, two-tailed, $P = 0.006$ and telomerase positive cell lines, $P = 0.014$ (Table 3.4).

3.3.8 Telomere mutation detection in the LoVo cell line

LoVo cell line is homozygous for a deletion in exons 4-8 of the *MSH2* gene (Wheeler *et al.*, 1999) and it shows MSI (Efstathiou *et al.*, 1999). The cell line is heterozygous in the Xp/Yp telomere adjacent sequence (haplotypes A and B) but the B-associated telomere was uninformative for telomere mutation detection. The haplotype A-associated telomere contained an interspersion of T and sequence variant repeats over the first eighty repeats (480 bp) and it was used for mutation detection.

In total, 93 copies of the Xp/Yp A-haplotype telomere were screened by TVR-PCR and seven mutant alleles were found (Figure 3.7 and Figure 3.8). Five mutations were deletions (clones 28, 33, 54, 60 and 64) of one to three telomere repeats and some of these deletion mutants were observed more than once among the 93 clones analysed. Another mutation was an insertion of five TTGGGG telomere repeats found approximately 50 repeats from the start of the telomere (clone 80). The last mutation could not be described

Vaco5 telomere maps

Xp/Yp haplotype C-associated telomere

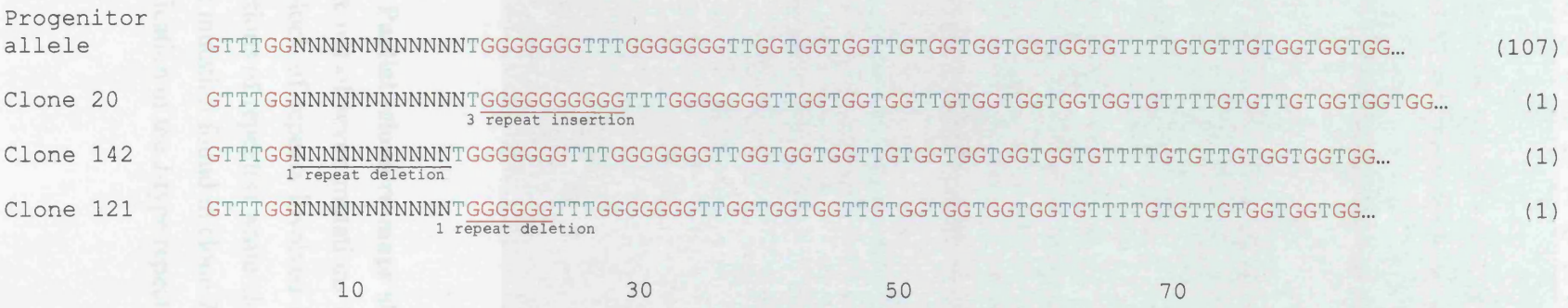


Figure 3.6 Comparison of the Xp/Yp associated telomere mutants from Vaco5 clones

Progenitor allele and clones with mutations in the haplotype C of the Xp/Yp associated telomere are shown. The block of repeats where the mutations are found is underlined and the number of telomere repeats involved in the mutation is given. In parenthesis at the end of each telomere map is shown the number of times the progenitor and the mutant alleles were observed. T, TTAGGG; G, TGAGGG; N, repeats that do not amplify with the primers used.

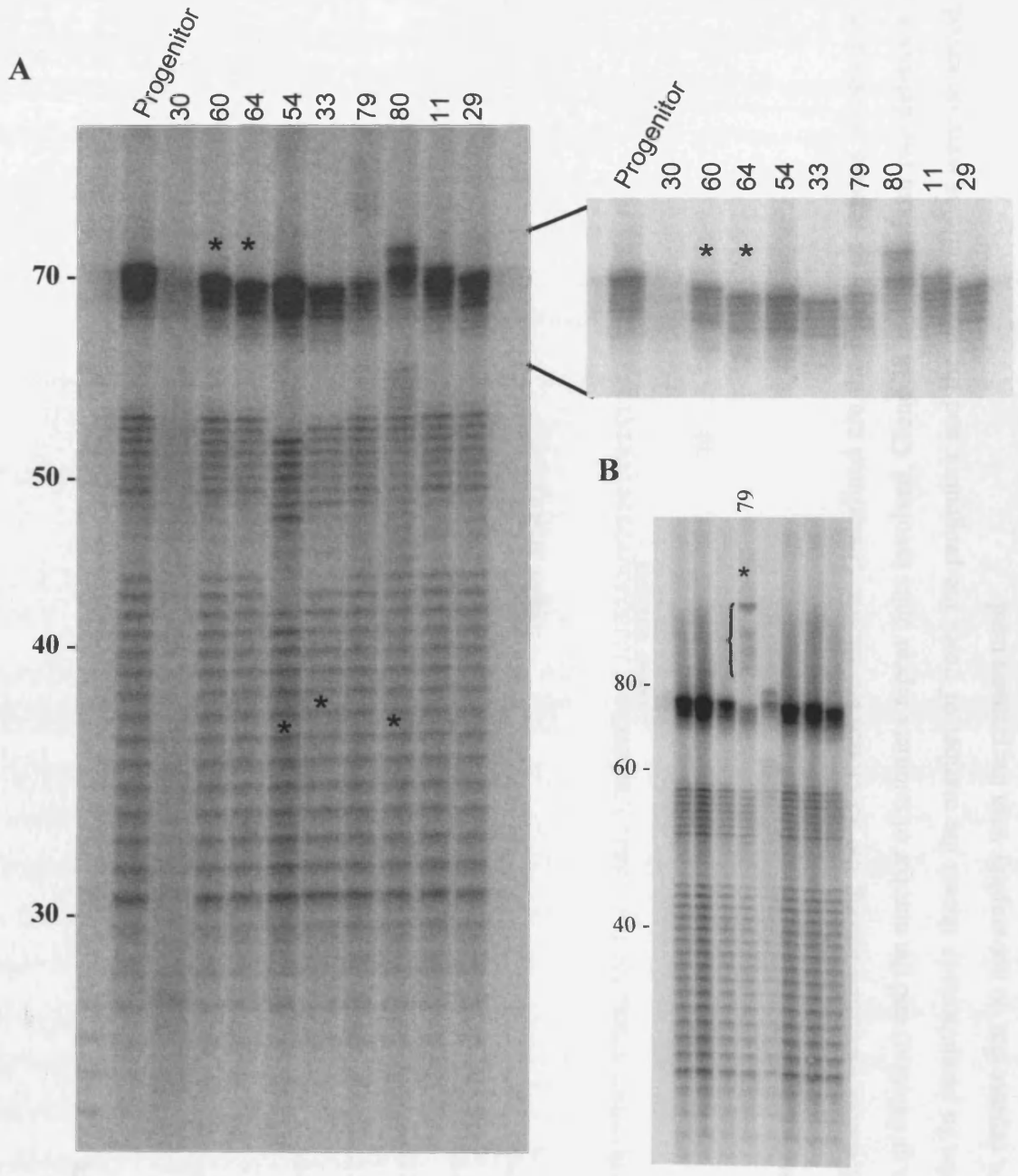


Figure 3.7 Partial telomere maps showing mutation in clones from the LoVo cell line. Six out of seven mutations found in LoVo cell line are shown. The (*) shows the block of repeats involved in the mutations. Clones 33, 54, 60 and 64 showed deletions of repeat(s) while clone 80 showed a five-repeat insertion. In (*B*) is shown the mutation found in clone 79 (indicated with bracket). The pictures only show amplification of the J-type repeat (TTGGGG).

LoVo telomere maps

Xp/Yp haplotype A-associated telomere

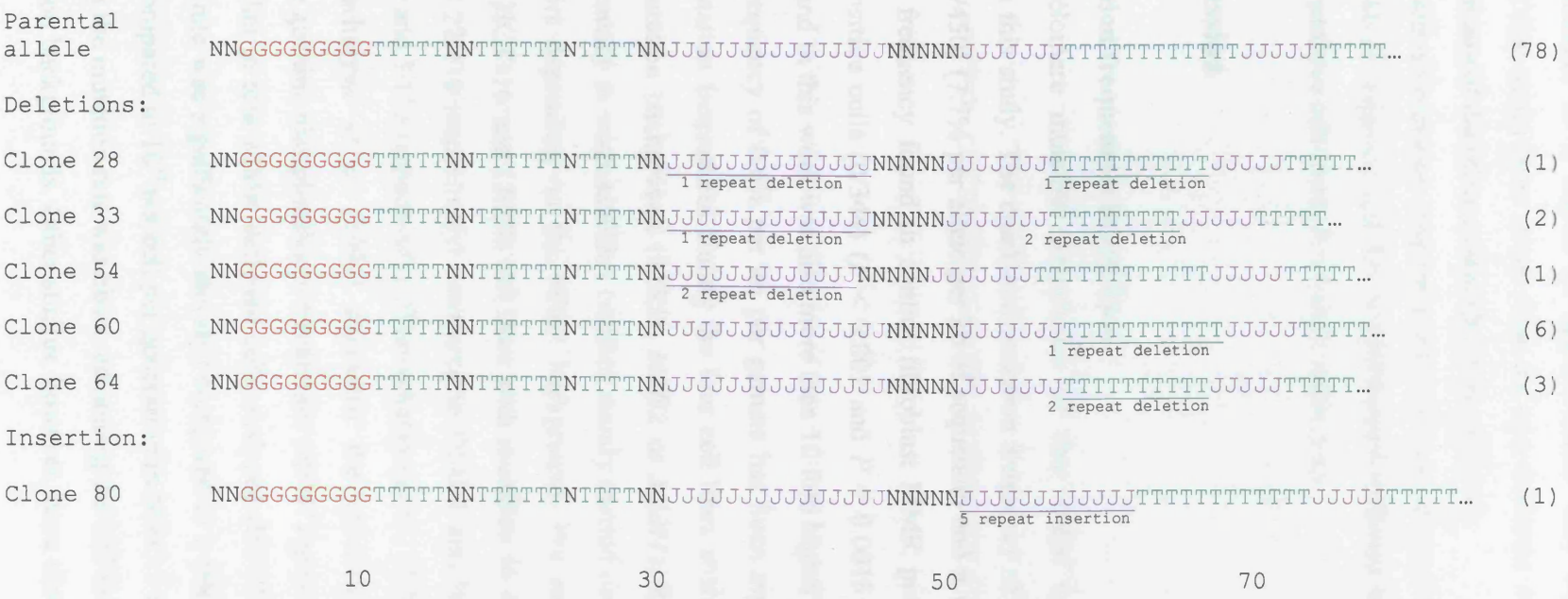


Figure 3.8 Comparison of the Xp/Yp associated telomere mutants from LoVo clones

Maps from 6 out of 7 mutant Xp/Yp telomeres identified among the 93 clones are shown. Underlined are the blocks of repeats involved in the mutation together with the type (insertion or deletion) and the number of telomere repeat units involved. Clone 28 and 33 harbour two deletions. All other clones show one mutation per allele. In parenthesis is shown the number of times the progenitor and the mutant alleles were observed. T, TTAGGG; G, TGAGGG; J, TTGGGG; N, repeats that do not amplify with the primers used.

definitively because the mutation was located more than one hundred telomere repeats into the repeat array (clone 79; Figure 3.7B). In this region, the resolution of the amplified products is poor thus making it difficult to read telomere maps. The mutation appears to be either a duplication of a region of approximately fifty repeats or a more complex rearrangement that could have arisen from an inter-telomere event. This type of mutation was not seen in any of the other mismatch repair cell lines.

The telomere mutation frequency was 7.5% per allele (95% CI, 3.5% - 15%) or 16% per allele per kb of sequence and this is significantly different from either primary cell lines or telomerase positive cell lines ($P < 0.001$; Table 3.4).

3.4 Discussion

3.4.1 Mutation frequencies in cell lines

The telomere mutation frequencies of four colon cancer cell lines have been determined in this study. The combined mutation frequency of the four cell lines was 3.6% per allele (16/450) (7.7% per allele per kb of sequence) and it was significantly higher than the mutation frequency found in control fibroblast MMR proficient cells (0/504) and in telomerase positive cells (0/344) ($P < 0.001$ and $P = 0.0015$ respectively). The mutation frequency found in this work was also more than 10 fold higher than in the germ-line, where a mutation frequency of 0.6% per kb per gamete has been reported (Pickett *et al.*, 2004). However, mutation frequencies among the four cell lines analysed were similar regardless the MMR mutation background (*MSH6*, *MSH2* or *MLH1*). Reports have shown that the mutation frequency at microsatellites (studies mainly carried out in mono- and di-nucleotide repeats) differs depending on the MMR background. For example a higher fraction of subclones of HCT116 and LS180 cell lines with mutation in the *MLH1* gene showed MSI (117/180 and 22/110 respectively) compared to DLD-1 and HCT15 cell lines defective in *MSH6* (8/130 and 3/113 respectively) (Bhattacharyya *et al.*, 1995; Ku *et al.*, 1999).

Bhattacharyya *et al* (1995) measured the spontaneous mutation rate of the hypoxanthine guanine phosphoribosyl transferase (*HPRT*) gene in MMR deficient cell lines. They found that in four MMR deficient cells analysed (HCT15, DLD-1, LS411, HCT116) the mutation rate was significantly elevated compared to a fibroblast control cell (mutation rate of 10^{-5} compared to 10^{-8} per cell per generation in MRC-5 control cell line), however no differences in the mutation rate was observed among the cell lines analysed regardless of the MMR mutation backgrounds. Other studies, however, have shown that the mutation rates at

the *HPRT* gene vary according to the MMR mutation background of the cells (Eshleman *et al.*, 1995; Branch *et al.*, 1995; Glaab and Tindall, 1997). For example, mutation rates in cell lines with *MLH1* mutations were approximately 10 times lower than cell lines with mutations in *MSH6* or *PMS2* ($4X10^{-6}$ and $2X10^{-5}$ mutations per cell per generation respectively; Glaab and Tindall, 1997). Thus, it is evident that not all loci are affected in the same way in cells with defects in MMR genes and therefore it may not be surprising to have found similar mutation frequencies at telomeres in cell lines with different MMR mutation backgrounds.

There are some factors that could influence the mutability of repeat sequences. It seems that the composition of the repeat array and the sequence surrounding the arrays could have an effect on the mutation dynamics of microsatellites. Zhang *et al* (2001) analyzed 29 mono-nucleotide repeats composed of $(A)_8$ or $(G)_8$ tracts on chromosome 22 in a region of ~1 Mb. It was found that $(G)_8$ tracts mutated more often (mutation frequency of 33.9%) than $(A)_8$ tracts (mutation frequency of 7.7%), however there was a substantial variation in the mutation frequency between $(A)_8$ tracts. This indicates that the position and the composition of the arrays are important for their stability. Furthermore, it was seen that homogeneity within microsatellites can increase their instability (Crawford *et al.*, 2000) and that interruptions in a homogeneous array can decrease the instability of microsatellites (Chong *et al.*, 1995). So, for example, unstable tracts of $(CAG)_{40}$ repeats can be stabilized by an interruption of a single CAT repeat (Chong *et al.*, 1995). Therefore, if these factors also affect telomeres, then the telomere mutation frequency reported in this work will be an underestimate of the mutation frequency further into the telomere which is composed of homogeneous TTAGGG repeats.

Mutations in tumour suppressor genes and oncogenes such as *TP53*, *APC*, *K-RAS*, β-catenein are proposed to have an effect on the genome stability of cancer cells and to promote tumour progression. The cell lines used in this work contained mutations in some of these genes (Table 3.1) and therefore, it is difficult to rule out the possibility that mutations in genes such as *TP53*, rather than the MMR genes, effected telomere stability. Nonetheless, the four cell lines used in this work do not share the same combination of mutations at tumour suppressor and proto-oncogenes, *e.g.*, LoVo cell line has mutations in *K-RAS* but not in *TP53* whilst Vaco5 has mutations in *TP53* but not in *K-RAS*. Given that the telomere mutation frequencies between the cell lines were similar, it could indicate that those genes do not have a direct effect on mutations at the start of the telomere repeat array.

3.4.2 Mutation frequencies in tumours

In Pickett *et al* (2004) the telomere stability was analysed in two panels of colon tumours. The telomere maps in the tumours were compared to telomere maps in a matched normal colon or blood sample and the mutation frequency was 5.8% per allele in a randomly collected panel of sporadic colon samples; however the mutation frequency increased to 18.6% (18/97) per allele in samples with instability at the polyA tract of the *TGFβRII* gene, used as an indicator of MSI. Furthermore, five of the tumours had somatic mutations in the *MSH2* gene and the mutation frequency per allele was 35% (7/20). The mutation frequency observed in this work (3.6%; 16/450) is significantly different to the observed in a panel of tumours that display MSI (Fisher's exact test, two-tailed; $P < 0.001$). It is not clear why there are such differences between the mutation frequency in the cell lines characterised here and the tumour samples analysed by Pickett *et al.* However Pickett *et al.* suggested that the mutations arose early in the development of the tumours because the tumour samples appeared to be dominated by one mutant allele. Therefore, it seems possible that the telomere mutations seen in the colon carcinomas originated before tumour formation or very early in tumour development but as tumorigenesis progressed the telomeres became more stable. Alternatively, as the MMR defective cell lines used in this study have been in culture for many generations, it is possible that they are better adapted to long term MMR deficiency and do not show a high telomere mutation frequency.

3.4.3 Mutation spectrum

Mutations were scored when a telomere map from a clone differed from the pattern observed in the parental allele in bulk genomic DNA. The number of loss and gain of repeats found in this study (8/16 and 3/16) was statistically similar to that observed previously in tumour samples (15/24 and 4/24) (Pickett *et al.*, 2004). Therefore the mutation process appears to be biased towards loss of repeats at telomeres. Monoucleotide microsatellites also show a bias towards loss of repeats in tumours (Parson *et al.*, 1995; Bacon *et al.*, 2001) and in colorectal cell lines defective in one of the MMR genes (Zhou *et al.*, 1997). For example, the mutation spectrum in eight mononucleotide repeat microsatellites was analysed in MSI cell lines and it was found that the majority of mutations were loss of repeats. The mutations also range in size going from losses of up to 16 repeat units compared to shorter gains of up to 3 repeat units (Zhou *et al* 1997).

The mutation spectrum of 122 autosomal tetra-nucleotide microsatellites has been investigated in families and it was found that the rate of deletions increases exponentially with allele size however insertions were constant in spite of allele size. Interestingly, the overall rate of deletions and insertions was equal, explaining the stationary allele distribution of microsatellite (Xu *et al.*, 2000). Therefore, it has been proposed that in the germ-line, long microsatellites favour contractions while short microsatellites tend to expand until they reach an equilibrium state. In this work, although there were significantly more deletions than insertions, the overall gain or loss of telomere repeats was equal since telomere insertions were bigger (involving 3 to 5 repeats) than deletions (involving 1 to 2 repeats). It is possible therefore that in the MMR defective cell lines analysed in this work there is not an overall change in telomere length produced by MMR defects.

The HCT15 cell line is homozygous for a mutation in the *MSH6* gene. *MSH6* is involved in the repair of base mismatches and short insertion/deletion loops. It has been reported that the HCT15 cell line displays mutation in mono- but not in di-nucletide microsatellites (Cottu *et al.*, 1996). In this work, four telomere mutations were found in HCT15, three of them were repeat-type changes while the other mutation consisted of an insertion of five telomere repeats. The insertion of five repeats is not consistent with previous studies of the mutation spectrum at mononucleotide repeats in this cell line or with the proposed role of *MSH6* however it is difficult to extrapolate what is seen in mono-nucleotides to telomeres. It is well known that the mutation frequencies at microsatellites changes between different loci, probably depending on their composition or their chromosome location. For example, a study of microsatellites and minisatellites stability in tumours from sporadic colon cancer and HNPCC patients showed that all samples with MSI were also unstable at minisatellites, however three tumours from HNPCC patients with mutations in one of the MMR genes, showed instability at minisatellites but not at microsatellites (Coleman *et al.*, 2001).

3.4.4 Mechanism of telomere mutations

All the mutations described so far (insertions, deletions and repeat-type changes) could be explained by intra-allelic events. It is proposed that the process that induces instability at repetitive sequences is generated by replication slippage (Schlotterer and Tautz, 1992) however recombination processes such as gene conversion and unequal sister-chromatid-exchange also originated in replication *e.g.* stalled replication forks, could explain larger insertions/deletions of repeats (Weber and Wong, 1993; Lundin *et al.*, 2002; Oda *et*

al., 2005). Two of the mutations found in this work were not consistent with simple loss or gain of repeats. One of them was found in the LS411 cell line (12-A haplotype) and comprised a change in the pattern of repeats of a block of 11 telomere repeats. This change did not affect the surrounding repeats so there was no overall loss or gain of repeats and it could have originated from more than one mutation event. Another mutation that could not be fully defined was found in LoVo cell line (*MSH2* deficient). This mutation appeared as the addition of three blocks of TTGGGG repeats (Figure 3.7) and comprised a region of probably 50 telomere repeats. It is possible that these additional blocks of repeats arose through an intra-allelic duplication event that might have arisen through an unequal sister chromatid exchange as the progenitor allele contains three blocks of TTGGGG repeats.

The MMR system inhibits homologous recombination of imperfectly matched sequences (Selva *et al.*, 1995), a process known as homeologous recombination, which may be elevated at yeast telomeres in the absence of MMR genes. Rizki and Lundblad (2001) showed that loss of mismatch repair proteins increase cellular proliferations in a double mutant strain of *Saccharomyces cerevisiae* defective for telomerase and one of the MMR genes. The increase survival in the absence of telomerase is though to be facilitated by telomeric recombination. Moreover, after inhibition of telomerase in a clone of the HCT15 cell line (*MSH6* deficient) an increase in telomere sister chromosome exchanges was observed (Bechter *et al.*, 2004). Therefore the telomere mutation found in LoVo cell line in this work might be the result of an inter-molecular event; however the poor resolution of the telomere map for this mutation does not allow the mutation mechanism to be determined definitively.

Chapter 4

Analysis of Telomere Stability in Colorectal Tissue Using a Single Molecule Approach

4.1 Introduction

4.1.1 Colon organisation

The intestinal epithelium is a self-renewing tissue which is formed by a single layer of epithelial cells divided by two main regions in the small intestine: a proliferative region or crypt and a differentiated region called the villus. In the colon, the crypt comprises almost all the intestinal epithelium and the villus is not present. The differentiated cells are located at the top third of the crypt, in the surface epithelium (Figure 4.1; Radtke and Clevers, 2005; Crosnier *et al.*, 2006). The intestinal epithelial stem cells are located at the bottom of the crypt. These cells renew themselves through life to produce transit-amplifying cells located in the remainder of the crypt. Transit-amplifying cells become differentiated three or four cell divisions later migrating upwards to the villus (small intestine) or to the surface epithelium (colon) to be shed finally into the lumen of the intestine (Bjerknes and Cheng 1999).

Intestinal stem cells produce four different cell types: absorptive, enteroendocrine, goblet and Paneth cells. Absorptive cells are the most common cell type in the small intestine; these cells secrete hydrolases and absorb nutrients. Enteroendocrine cells represent less than 1% of all cells in the intestine and secrete various types of gut hormones. The goblet cells are the most common cells in the large intestine and its function is to secrete protective mucus. The Paneth cells are located at the bottom of the crypt and are only found in the small intestine. These cells secrete antibacterial proteins (such as lysozyme, defensins and cryptdins) to keep the microbial content in the intestine stable. All four cell types are produced by a single stem cell, thus the adult crypt has a monoclonal origin (Kirkland, 1988).

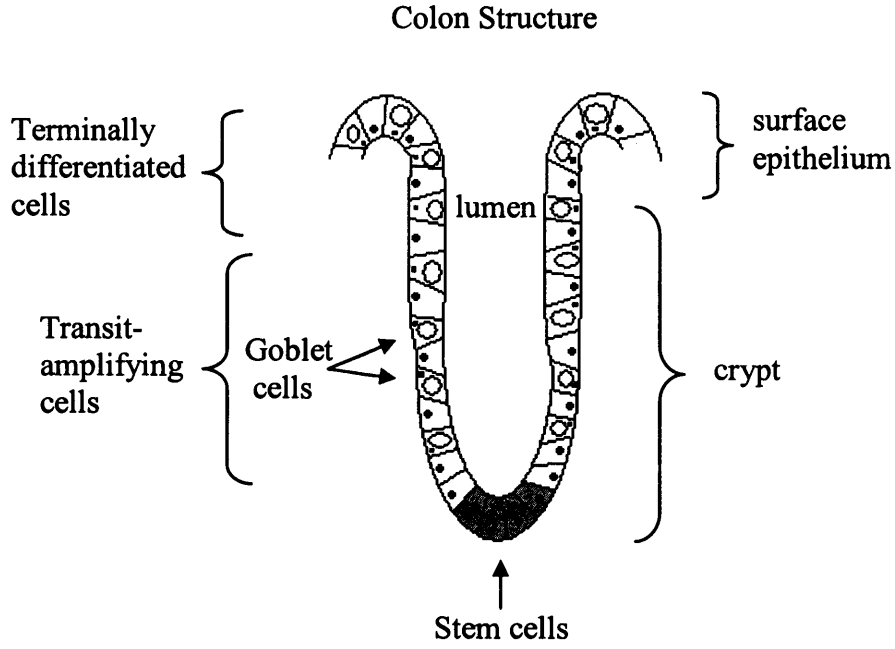


Figure 4.1 Structural organisation of the large intestine. The stem cells are located at the bottom of the crypt. Differentiating cells migrate upwards until they reach the surface epithelium. Cells are shed into the lumen of the intestine and replaced constantly.

4.1.2 Regulation of cell proliferation in the colon

The main pathway that regulates the colon stem cell function and proliferation is the Wnt signalling pathway. The *WNT* genes encode signalling proteins that control the fate and proliferation of epithelial cells. In humans, the *WNT* family consist of 19 genes that bind to *FZ* (frizzled) receptor family (Reya and Clevers, 2005). The binding of Wnt to its receptor, activates a cascade of events that leads to the accumulation of β-catenin in the nucleus which then binds to its receptors Lef or Tcf4 to promote activation of target genes that lead to cell proliferation. However in the absence of Wnt proteins, β-catenin interacts with a complex of proteins called the destruction complex formed by adenomatous polyposis coli (APC), glycogen synthase kinase 3β (GSK-3β) and axin. The interaction of β-catenin with the destruction complex leads to its phosphorylation by GSK3β and finally to proteosome-dependent degradation (Aberle *et al.*, 1997). Using RNA probes against members of the *WNT* and *FZ* family, Gregorieff *et al.* (2005) demonstrated the presence of several members of both families in the crypt of the small intestine and the colon proving the presence of this pathway in the generation of colon epithelium.

4.1.3 Colorectal cancer (CRC)

It has been proposed that serial mutations of oncogenes combined with inactivation of tumour suppressor genes causes CRC. For example, some of the alterations found in CRC are mutations that activate the *K-RAS* proto-oncogene and allelic losses containing tumour suppressor genes such as *APC* and *TP53*. It has been proposed that the accumulation of mutations rather than the order of events is important to lead to CRC (Vogelstein *et al.*, 1988). CRC can occur sporadically by inactivation of gatekeeper genes (mainly by modifications of gene promoters), however hereditary forms of CRC have also been described and account for approximately 10% of all CRC (Lynch and de la Chapelle, 2003). These tumours are divided into two main groups accordingly to the presence or absence of polyps: familial adenomatous polyposis (FAP) and hereditary nonpolyposis CRC (HNPCC).

CRC is initially identified as an aberrant crypt focus, with different patterns of β-catenin staining. Two main models for early neoplasia at the colon have been proposed. The first suggests that the mutant cells are found at the surface epithelium and from there they migrate laterally and downwards to form new crypts (Shih *et al.*, 2001). These conclusions were obtained by analysing 35 adenomas for *APC* mutations and changes in β-catenin expression and localisation as a marker for Wnt pathway activation. In all cases mutations at the *APC* gene and nuclear localisation of β-catenin were located at the surface epithelium but not at the base of the crypt where cells appeared to be morphologically normal, thus it was concluded that malignant cells are located in the surface of the crypt.

Since colonic stem cells are located in the base of the crypt, the second model for colorectal cancer formation suggests that mutations/inactivations that lead to cancer originate in the stem cells and then spread upwards to occupy the rest of the crypt (Preston *et al.*, 2003). This model was also proposed on the basis of β-catenin staining in adenomas and tumours from FAP patients. It was found that the dysplastic cells where located in the base of the crypt towards the top where β-catenin staining was hardly found. In addition, they observed that the mutated clone expands not by lateral migration but via crypt fission in early stages of tumour formation. However staining of β-catenin was also observed in the surface epithelium of bigger-size adenomas, therefore it was concluded that spread via the surface may appear in later stages of tumourigenesis.

4.1.3.1 Familial adenomatous polyposis (FAP)

Patients that show the FAP syndrome have numerous adenomatous polyps in their colon and a high probability of developing CRC by the age of 40 to 50 years (Nieuwenhuis and Vasen, 2007). FAP patients inherit one mutant copy of the *APC* gene (Groden *et al.*,

1991) and will develop cancer when a second hit in the normal allele is acquired. *APC* is classified as a tumour suppressor gene involved in transcription and cell cycle control. Mutations in this gene are not limited to FAP patients, they also occur in somatic cells leading to most of the sporadic CRC and are identified by analysing the accumulation of β-catenin protein in the cytoplasm and nucleus. Accumulation of β-catenin in the nucleus induces the transcriptional activity of genes involved in cell proliferation. Most mutations in the *APC* gene found in FAP and sporadic colon cancers occur in the β-catenin binding domain thus leading to a deregulation of the Wnt pathway and accumulation of β-catenin in the nucleus. Alternatively, some colon cancers, where no mutations in the *APC* gene have been found, contain mutations in the phosphorylation N-terminal regulatory domain of β-catenin that prevent its degradation by Apc and thus promoting its stability and accumulation in the nucleus (Sparks *et al.*, 1998).

In addition, mutations in the *APC* gene lead to defects in mitosis; specifically it contributes to defects in mitotic spindle formation leading to some of the aberrations observed in colorectal cancers such as high levels of aneuploidy. Furthermore, depletion of *APC* reduces the levels of apoptosis that together with defects in mitotic spindle formation lead to an increase in chromosome instability. These observations have been confirmed in a MSI colon cancer cell line (HCT116) with no characterized mutations in the *APC* gene by either transfecting the cells with a siRNA (Dikovskaya *et al.*, 2007) or by expressing truncated versions of the *APC* gene (Tighe *et al.*, 2004). Studies in mice have shown that deletions of both alleles of the *APC* gene cause rapid formation of colorectal adenomas (within 4 weeks) however the lack of only one *APC* functional gene did not cause adenoma formation (Shibata *et al.*, 1997).

4.1.3.2 Hereditary colorectal cancer (HNPCC)

Similar to FAP patients, HNPPC patients display an early onset of colorectal cancer, with an average age of 45 years. HNPCC is generated by germ-line mutations in one copy of a MMR gene. During tumourigenesis, the unaffected MMR gene is lost, mutated or silenced leading to accumulation of mutations throughout the genome and eventually to CRC formation. The majority (~90%) of mutations found in MMR genes are present in the *MLH1* and *MSH2* genes. However, mutations in *MSH6*, *PMS2* and *PMS1* genes have also been described (Nicolaides *et al.*, 1994; Miyaki *et al.*, 1997; Hendriks *et al.*, 2006).

One of the hallmarks of HNPPC syndrome is an increase in the instability of short tandem repeats known as microsatellite instability (MSI). Most microsatellites are found in non-coding sequences where MSI might not have an impact in the development of cancer.

However it is proposed that MMR defects can contribute to the generation of CRC by promoting mutations in some of the genes involved in CRC formation that contain coding microsatellites such as the *APC* and *K-RAS* genes. There are other genes with coding microsatellites involved in cell proliferation and apoptosis that mutate at a high rate in patients with MMR defects. For example, the transforming growth factor β type II receptor (*TGFβRII*) which contains a tract of 10 adenines shows frameshift mutations in approximately 90% of colorectal cancers with MSI (Parsons *et al*, 1995). Other genes that contain repeat arrays within their coding regions and that are affected by MMR defects are the apoptosis-promoter gene *BAX*, which mutate in about 40% of CRC cases with defects in MMR (Rampino *et al.*, 1997). In the same way, MSI affects sequences found in other genes important to tumourigenesis such as the insulin-like growth factor II receptor (*IGFIIR*) and other MMR genes like *MSH6* and *MSH3* (Souza *et al.*, 1996; Malkhosyan *et al.*, 1996).

4.2 Aims

Mutation analysis at telomeres in a panel of colorectal cancer samples have shown that telomeres mutate *in vivo* and those mutations were strongly correlated with the presence of MSI (mutation frequency of 18.6%) and mutations in the *MSH2* gene (mutation frequency of 35% per allele). Telomere mutations were determined by comparing telomere maps carried out by TRV-PCR using bulk DNA from either blood or normal colon to tumour DNA (Pickett *et al.*, 2004). TVR-PCR on genomic DNA from tissue gives information of the most common allele in population of cells contributing to the tissue. The mutations observed in colon cancers (Pickett *et al.*, 2004) appeared to be intra-allelic and if the mutation process is dependent on replication then it is possible that tumours will contain multiple rare telomere mutations that arise later during tumour proliferation. Therefore one aim was to develop a single molecule PCR approach for telomere mutation detection in tissue samples. This technique was then used to assess telomere stability in normal colon tissue and colon cancers.

4.3 Results

4.3.1 Telomere Variant Repeat mapping from single telomeres

4.3.1.1 Single telomere length analysis (STELA) optimisation

In order to analyse telomere stability within tumours, a single molecule approach to detect telomere mutations was developed. This technique consisted of the amplification of full length single telomeres from single molecules by STELA. Each product was then analysed by TVR-PCR to determine the interspersion of TTAGGG repeats with sequence-variants for mutation detection by comparison between molecules.

The STELA technique was used to amplify full length of telomeres. This technique was described in Baird *et al.* (2003) and consists of digesting high molecular DNA with an enzyme that does not cut telomere sequences (*e.g. EcoR*I) and ligating an oligonucleotide (telorette) to the telomere end (Figure 4.2). The telorette consists of 7 nucleotides complementary to the 3' overhang of the telomere and 20 nucleotides of unique sequence. Once the telorette has annealed, it is ligated to the 5' end of the C-rich strand (Figure 4.2). After that, long-range PCR is carried out using a primer (teltail) that shares its sequence with the 5' end of the telorette oligo and a chromosome-specific or an allele-specific primer located in the subtelomeric region.

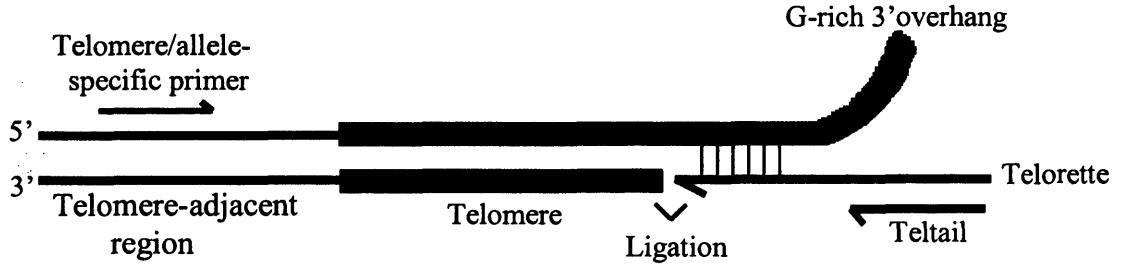


Figure 4.2 STELA technique. A telorette oligonucleotide is annealed to the G-rich overhang and ligated to the 5' end of the C-rich strand. PCR is then carried out using a chromosome or allele-specific primer that anneals in the telomere-adjacent region together with the teltail primer. Teltail is the same sequence as the 20-bp tail of telorette.

In the original description of the STELA method, it was proposed that the ligation of the telorette to the 5' end of the C-rich strand was essential. The ligation step was carried out at 35°C for 12 hrs using T4 DNA ligase (1 unit), ligation buffer, 0.9 μM of telorette and 10 ng of DNA in a 10 μl volume. However, while establishing the method for this study further controls were performed and it was found that surprisingly the ligation step is not required. DNA was digested and annealed to the telorette and then incubated with or without ligase. Subsequent PCR with the teltail and telomere adjacent primer generated a similar number of products with similar size distribution irrespective of whether a ligation step was included (Figure 4.3).

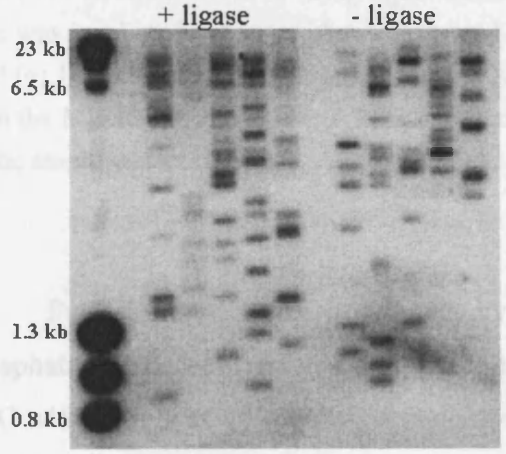


Figure 4.3 STELA at the Xp/Yp telomere using DNA from the MRC-9 cell line. STELA was carried out with DNA incubated with (+) and without (-) ligase.

DNA ligase is an enzyme that catalyses the formation of phosphodiester bonds between adjacent 3'-hydroxyl and 5'-phospate groups. It can ligate double-stranded cohesive DNA termini or nicks and it is ATP and Mg^{2+} dependent. In order to determine whether lack of ATP would stop the generation of STELA products, DNA was incubated with and without ligase as described above and ATP was not added to the reactions (Figure 4.4). The results showed that ATP was not required to generate STELA products, however there were fewer bands in the reaction where neither ligase nor ATP was added (Figure 4.4). Thus it is possible that ATP had an effect on the amplification. Further experiments where the ligase step was omitted and ligation buffer was not added to the reaction (so ATP was absent) showed that the number of bands and the intensity of them were similar between the reactions where the ligation step was included. Thus, it was concluded that ATP does not have an effect on the amplification of full length telomeres.

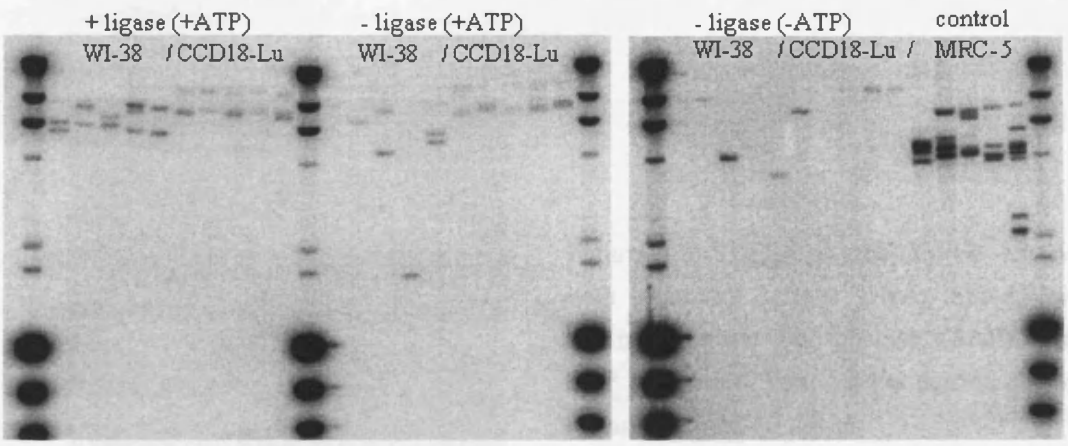


Figure 4.4 STELA at Xp/Yp using ligase buffer with or without ATP. DNA from two different cell lines was incubated with or without ligase and these reactions were performed in ligase buffer with ATP (+; 1 mM final concentration) or without ATP (-). As a control, STELA was conducted on DNA from the MRC5 cell line that had been digested and ligated to the telorette in ligase buffer provided by the manufacturer (USB).

Further tests were performed by incubating the DNA with shrimp alkaline phosphatase (SAP) to remove the 5'-phosphate group. The DNA was incubated for 15 min at 37°C with 2 units of SAP and then the enzyme was inactivated by heat treatment at 65°C for 15 minutes (Figure 4.5). Samples treated with SAP and then incubated with or without ligase generated STELA products that range from ~1 to 9 kb.

If the telorette has annealed to the G-rich overhang but not ligated to the C-strand then hot-start PCR should prevent the formation of STELA products. Therefore, hot-start PCR was carried out by adding the DNA polymerase mixture (*Taq*:*Pwo* 10:1) after the first PCR denaturation step (96°C for 30 sec). As expected the hot-start PCR was enough to prevent the generation of STELA products when the telorette was not ligated to the C-terminal end of the telomere (Figure 4.6). Since the distribution of bands produced by STELA performed with or without ligase is the same, all subsequent STELA reactions were carried out without incubating the DNA with ligase. This is the main modification made to the technique described in Baird *et al* (2003).

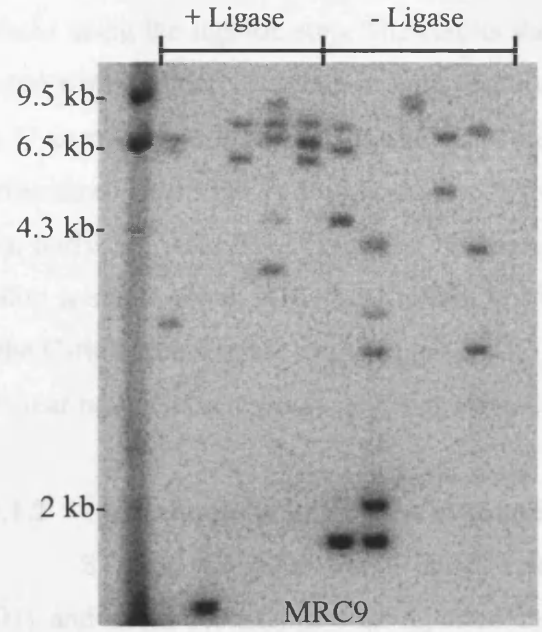


Figure 4.5 STELA at the Xp/Yp telomere with digested MRC9 DNA, treated with shrimp alkaline phosphatase (SAP). After SAP incubation, ligation was carried out with (+) and without (-) ligase.

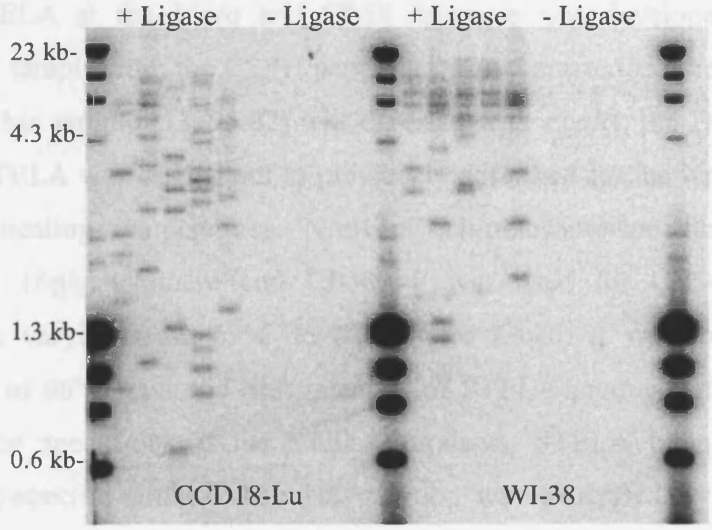


Figure 4.6 Hot-start STELA at the Xp/Yp telomere. Two different DNAs from cell lines CCD18-Lu and WI-38 were incubated in the presence of telorette with (+) or without (-) ligase. During the STELA-PCR, *Taq:Pwo* polymerase mixture was added once the PCR reactions reach 96°C during the first PCR cycle.

Additional control tests were performed in order to amplify Xp/Yp telomeres by STELA. For example, the DNA from MCR9 cell line was incubated with 40 units of mung-

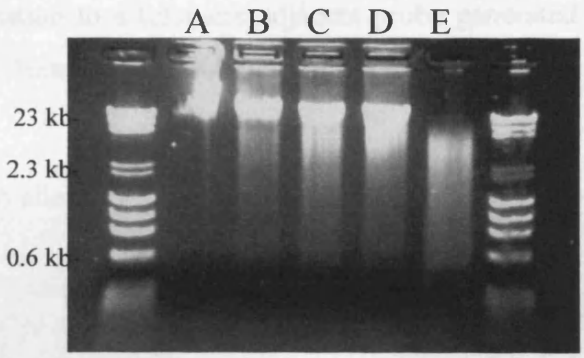
bean nuclease (MBN) in order to degrade the 3' overhang of the telomeres either with or without using the ligation step. The results showed no amplification of telomeres in samples treated with the MBN enzyme in both reactions (with or without ligation step), proving that the 3' overhang is essential for the annealing of telorette and for telomere amplification. Furthermore, the STELA technique also proved to be reliable when amplifying telomeres from fairly degraded DNA (Figure 4.7). Thus, it allows the study of material where the DNA quality is sub-optimal. With these results it was possible to conclude that ligation of telorette to the C-rich strand is not required to amplify full length telomeres by STELA however it is not clear how telomere products are generated without ligation of the telorette.

4.3.1.2 Development of STELA at the 16p/q, 12q and CB54 telomeres

STELA technique was originally described at the Xp/Yp telomere (Baird *et al.*, 2003) and more recently it was reported for telomeres at 2p, 11q, 12q and 17p (Britt-Compton *et al.*, 2006). In this work, STELA was developed at the 16p/q (Coleman *et al.*, 1999) and CB54 (Dr Nicola Royle unpublished data) polymorphic telomeres. These telomeres are found in about 6% and 11% in the Caucasian population respectively.

STELA at the 16p/q and CB54 telomere was developed using DNA from a lymphoblast sample from the CEPH panel that was characterised previously to contain both telomeres. This sample (13293.02) was digested with *EcoR*I, the DNA was diluted to 250 pg/µl and STELA was carried out as previously described for the Xp/Yp telomere but using different annealing temperatures. Nitu14eC chromosome-specific primer was used to amplify the 16p/q telomere and CB38A-F was used for CB54 telomere. Annealing temperatures varying from 63°C to 67°C were tested. It was found that an annealing temperature of 66°C gave the best intensity of STELA products at both ends. In order to determine the specificity of the STELA products, STELA blots were hybridised with chromosome-specific probes. The 16p/q probe was generated by PCR amplification of genomic DNA with primers Nitu14eA and Nitu14eB that amplify a 178 bp fragment adjacent to the telomere, while the CB54 probe was amplified using primers CB38A-F and CB38A-R that amplify a 154 bp fragment from sequences adjacent to the telomere.

STELA at the 12q telomere was carried out using the chromosome-specific primer 12q-STELA with an annealing temperature of 66°C. Allele-specific 12q primers where not used in this work. A proportion (6%) of 12q telomeres is associated with a deletion in the telomere- adjacent sequence, the 12q deletion allele (12qΔ) and so STELA was developed



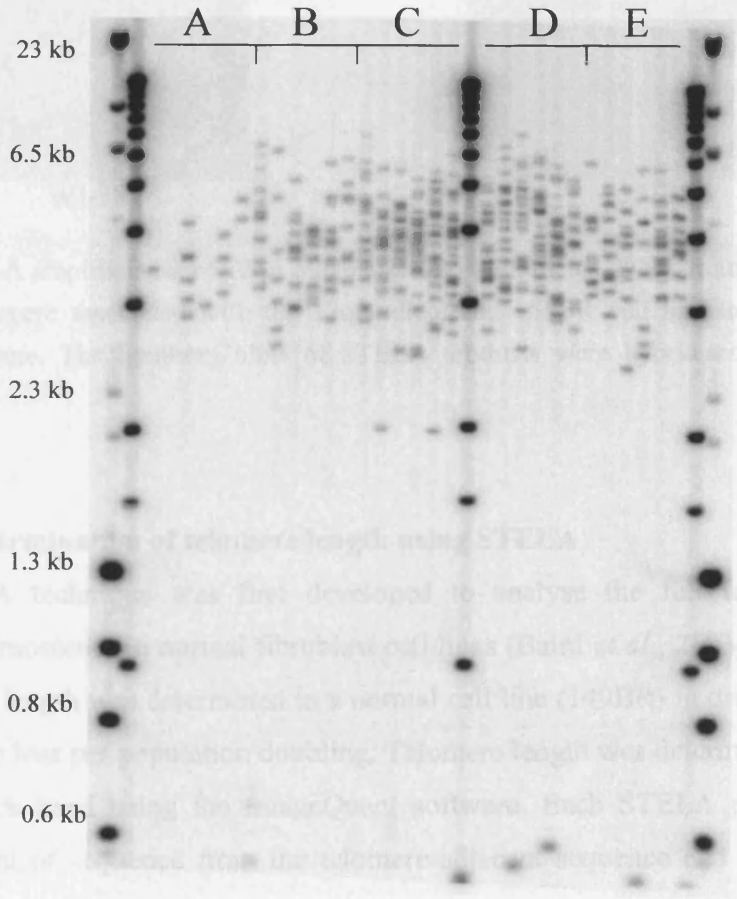


Figure 4.7 STELA at the Xp/Yp telomere of HT1080 cell line. High molecular weight DNA from HT1080 cell line (A) was degraded by sonication to different extents (B, C, D and E) and the partially-degraded DNA was used to carried out STELA at the Xp/Yp telomere without using the ligation step.

for these alleles using a telomere-adjacent primer (12q-null3) an annealing temperature of 65°C. Confirmation that these 12qΔ-STELA products were chromosome specific was achieved by hybridisation to a telomere adjacent probe generated using primers 12qB and pKSRV2 (Figure 4.8; Baird *et al.*, 2000).

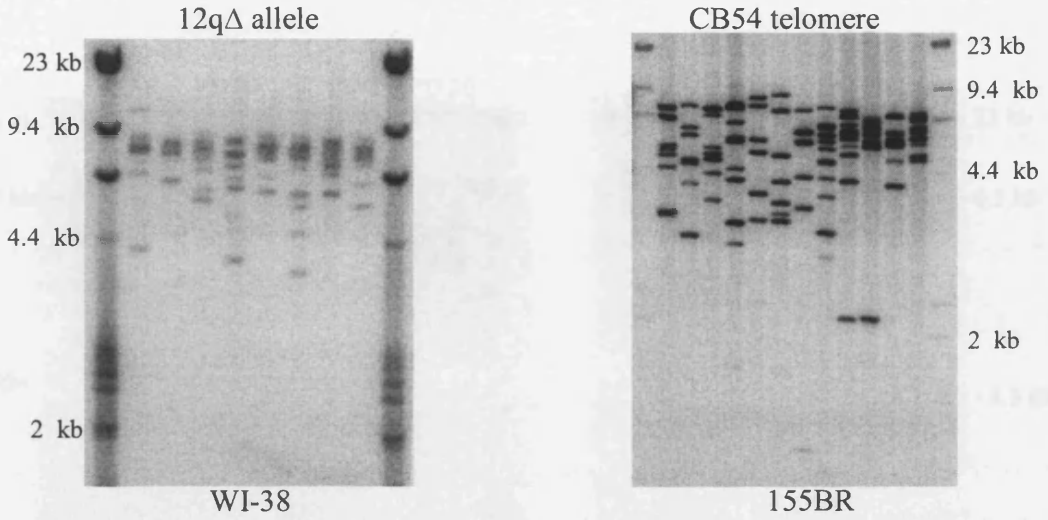


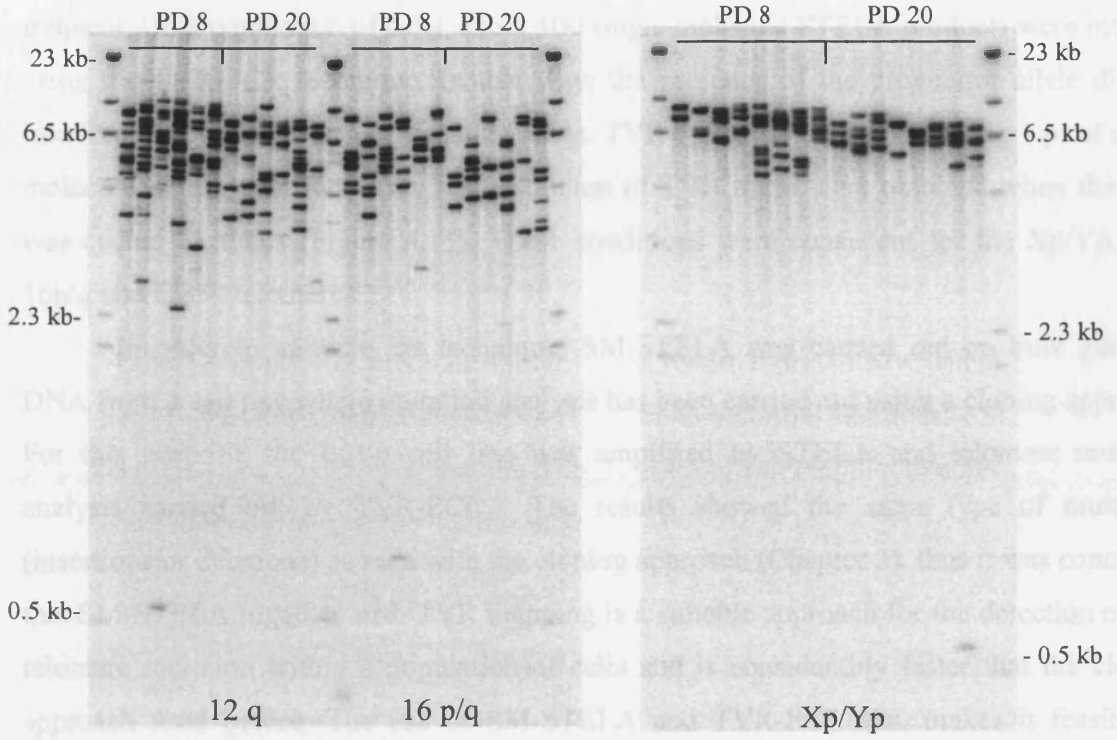
Figure 4.8 STELA amplification of 12qΔ allele and CB54 telomeres. DNA from WI-38 was used to amplify the telomere associated with the 12qΔ allele and 155BR cell line to amplify the CB54 associated telomere. The Southern blots of STELA products were hybridised with chromosome specific probes.

4.3.1.3 Determination of telomere length using STELA

STELA technique was first developed to analyse the full telomere length of individual chromosomes in normal fibroblast cell lines (Baird *et al.*, 2003). In this work, the mean telomere length was determined in a normal cell line (149BR) in order to establish the rate of telomere loss per population doubling. Telomere length was determined by measuring the size of each band using the ImageQuant software. Each STELA product includes a constant amount of sequence from the telomere-adjacent sequence and this was deducted prior to the calculation of the mean telomere length (406 bp of flanking sequence for the Xp/Yp, 544 bp for the 12q and 158 bp for the 16p/q were deducted).

In Figure 4.9 an example of STELA amplification at the Xp/Yp, 12q and 16p/q telomere from a normal cell line (149BR) is shown at two different population doublings. The mean and median telomere length of the Xp/Yp, 12q and 16p/q telomeres analysed was approximately 1 kb different between the longer telomere (Xp/Yp) and the shorter (16p/q) telomere. The rate of telomere shortening was calculated and all telomeres shortened at a

rate of 41 to 52 bp per PD which is similar to that reported previously, where a loss of 86 ± 42 bp per PD was described at the Xp/Yp telomere in a normal fibroblast cell line (Baird *et al.*, 2003). These results showed that the amplification of full length telomeres using STELA is not affected by the lack of the ligation step.



	12 q		16 p/q		Xp/Yp	
	PD 8	PD 20	PD 8	PD 20	PD 8	PD 20
Mean	5.97	5.38	5.44	5.14	6.5	6.07
Median	5.46	5.27	5.62	4.99	6.55	6
St Dev	2.30	1.94	1.28	1.32	1.05	0.64
St Error	0.29	0.31	0.22	0.24	0.16	0.11

Figure 4.9 STELA performed on the 149BR normal cell line using chromosome-specific primers at the Xp/Yp, 12q and 16p/q telomeres. STELA was carried out on DNA from two different PDs and the telomere length was determined using the ImageQuant software. Telomere length measures are shown in kb.

4.3.1.4 Detection of telomere mutations using single molecule STELA (SM-STELA) and TVR-PCR

In order to amplify single telomeres with the STELA technique, the DNA was diluted to about 30 to 50 pg per reaction using the conditions described in 2.2.7.3. All dilutions were made using tRNA as a carrier at a final concentration of 1 ng/µl. An example is shown in Figure 4.10. Half of the total PCR reaction was loaded into the gel (10 µl) reserving the other half for further analysis. To detect mutations at telomeres that arise at low frequency in a population of cells, about 100 single molecule STELA products were mapped using the TVR-PCR technique. In that way, the presence of the progenitor allele did not interfere with the detection of a mutant allele. TVR-PCR was carried out using 1 µl of single molecule STELA product. The best resolution of TVR maps were obtained when the PCR was cycled 21 times (Figure 4.11). These conditions were consistent for the Xp/Yp, 12q, 16p/q and CB54 telomeres.

In order to validate the technique, SM-STELA was carried out on bulk genomic DNA from a cell line where mutation analysis has been carried out using a cloning approach. For this purpose, the LoVo cell line was amplified by STELA and telomere mutation analysis carried out by TVR-PCR. The results showed the same type of mutations (insertions or deletions) as seen with the cloning approach (Chapter 3), thus it was concluded that SM-STELA together with TVR mapping is a suitable approach for the detection of rare telomere mutation within a population of cells and is considerably faster that the cloning approach used before. The use of SM-STELA and TVR-PCR also makes it feasible to determine the telomere mutation frequency and profile in tissue samples from patients.

4.3.2 Investigation of telomere stability in normal colon and carcinomas

4.3.2.1 Analysis of telomere stability in normal colon tissue samples

Stability at microsatellites has been studied in normal somatic cells using small-pool PCR approaches. Increase in MSI was seen in peripheral blood lymphocyte samples from 60-70 years-old donors compared to samples from 20-30 year-old donors (Coolbaugh-Murphy *et al.*, 2005). Therefore MSI in some normal somatic cells increases with age probably due to replication or repair defects. It has also been reported that normal colon tissue from patients carrying a mutation in one of the MMR genes, displayed low levels of MSI even before tumour diagnosis (Parsons *et al.*, 1995 and Alazzouzi *et al.*, 2005). These levels of instability were only detected using small-pool PCR approaches and not with conventional large scale PCR due to the low percentage of cells with mutant alleles.

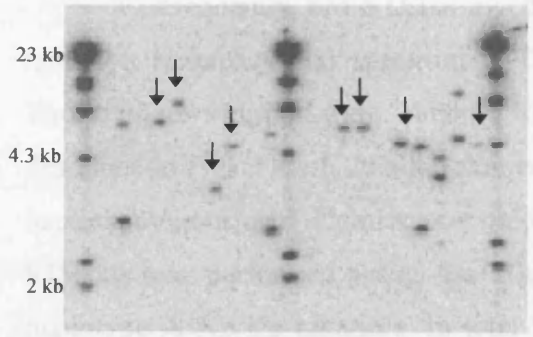


Figure 4.10 Single molecule STELA at the Xp/Yp telomere. *EcoR*I digested DNA from Vaco5 cell line was diluted to 50 pg per reaction to obtain full length telomere amplifications from single molecules as indicated by the arrows.

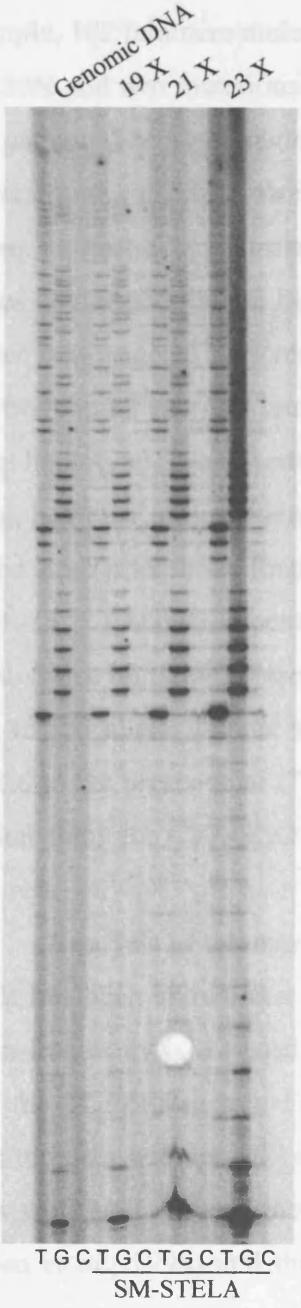


Figure 4.11 TVR at the haplotype C-associated Xp/Yp telomere from SM-STELA products from the Vaco5 cell line. TVR from genomic DNA and SM-STELA (using allele-specific primers) were carried out using 100 ng or 1 µl of STELA products respectively. The TVR-PCR was cycled 19 to 23 times and 20-21 cycles gave the best resolution.

Therefore, it is also possible that mutations at telomeres exist in normal tissues and that the age of the individual influences the frequency of such mutations.

Consequently, SM-STELA together with TVR-PCR was used to investigate telomere mutation frequency and spectrum in DNA extracted from normal colon tissue samples. These tissue samples were obtained from patients with sporadic colon cancer and are described in Picket *et al* (2004). Telomere mutation analysis was carried out at the Xp/Yp (A haplotype) associated telomere on a colon sample (2N) from a 66 year-old male patient. SM-STELA was performed using the allele-specific primer XpYp-415C that amplifies the haplotype A Xp/Yp telomere. In total 92 telomere molecules were screened for telomere mutations, but none were found. The same analysis was carried out in others colon samples, for example, 102 telomere molecules at the haplotype B Xp/Yp telomere were screened from sample 20N and two mutations were found. This sample was obtained from an 81-years old female patient. The two mutations found in this sample involved a block of CTAGGG variant telomere repeats. However, it was already known that CTAGGG-containing Xp/Yp telomeres are particularly unstable in the male germ-line (Hills, 2004) thus the two mutant telomeres identified here are likely related to CTAGGG repeat instability and not necessary associated with age. The screen for telomere mutations was expanded to another colon sample from an 88 year-old female individual (23N). Altogether 86 telomere molecules from the 12q haplotype B-associated telomeres were screened using SM-STELA and TVR mapping, however no mutations were found. Telomere mutation detection was also carried out at the Xp/Yp telomere (haplotype A) of the same sample (23N). This allele contains an array of 13 CTAGGG repeats and 6 mutations were identified out of 108 molecules screened. Since mutations were only present in the allele where CTAGGG repeats are found but not in the allele without such repeats, it was assumed that the mutations seen were associated to the presence of CTAGGG repeats and not necessarily to age. Further work and discussion about the CTAGGG repeat instability is presented in Chapter 6.

4.3.2.2 Analysis of telomere stability in colon carcinomas

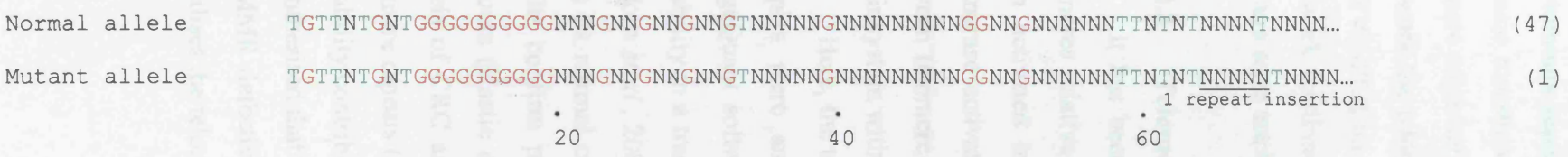
It has been shown that telomere mutations occur *in vivo* in colon tumours and the mutation frequency is elevated in tumours with MSI (assayed by the stability at the polyA tract of the *TGFβRII* gene). For instance, the telomere mutation frequency was 6.2% per allele in tumours without MSI compared to 18.6% per allele in tumours with MSI. From the tumours with MSI, five of them contained characterised mutations within the *MSH2* gene (Borrosen *et al.*, 1995) and their telomere mutation rose to 35% per allele (Pickett *et al.*, 2004).

In this work, two of the tumours with mutations at the *MSH2* gene (955 and 1314) were analysed by SM-STELA and TVR-PCR to investigate telomere stability at the Xp/Yp telomere. Mutations at the Xp/Yp telomere have been described in these tumours by comparing telomere maps from bulk DNA (from the colon carcinoma sample) to blood DNA from the same patients (Pickett *et al.*, 2004). The mutations described in Pickett *et al.* were observed in bulk DNA, so the majority of cells contained the mutation. That means it arose early in tumourigenesis or even before tumour formation. The purpose of this analysis was to examine whether the instability detected in these tumours is an ongoing process. If that were the case, mutations that arise at a low frequency would be detected by the single molecule PCR approach described above.

Tumour 955 is heterozygous for haplotype B and C in the flanking region of the Xp/Yp telomere. However, haplotype B and C are homozygous at the -30 SNP, where the allele-specific primer anneals for TVR-RCR, hence two different maps were generated by using the allele specific primer TS-30A in the TVR-PCR. The maps obtained were classified as haplotype B or C by comparing their telomere variant repeat distribution with published population data (Baird *et al.*, 1995). In total, 98 molecules were screened for mutations, 50 were the B haplotype-associated telomere and 48 the C haplotype-associated telomere. Analysis of bulk DNA from this tumour had identified a mutation in the haplotype B-associated allele and by SM-STELA and TVR-PCR it was found that this mutant allele was present in 76% of the haplotype B molecules screened, but no additional low frequency mutations were detected. Among the molecules analysed, 24% contained the same telomere map as seen in the normal blood DNA sample and this suggests that either they arose from normal cells within the tumour or the mutation arose very early during tumour formation. However one new mutation was identified among the 48 haplotype C-associated Xp/Yp telomere molecules analysed. The mutation frequency was 1% per allele (1/98; 95% CI, <0.001% - 6.1%) or 2.3% per allele per kb of sequence. This mutation consisted of the insertion of 1 telomere repeat about 65 repeats from the start of the telomere (Figure 4.12*A*).

Tumour 1314 is heterozygous at the telomere flanking region displaying haplotypes A and B. A total of 75 molecules were analysed at the haplotype B and it was found that 85% contain a mutation already described in Picket *et al.* (2004) and shown in Figure 4.12*B*. Among the 68 molecules with the mutated telomere map an additional new mutation was found and consisted of a deletion of one N-type telomere repeat and this mutation was presented in 4.7% of molecules screened (this mutation was observed three times). The mutation frequency was 1.3% (1/75) per allele (95% CI, < 0. 001% - 7.9%) or 2.2% per allele per kb of sequence.

A) Telomere Maps of 955 tumour sample. Xp/Yp telomere (Flanking haplotype C)



B) Telomere Maps of 1314 tumour sample. Xp/Yp telomere (Flanking haplotype B)

Figure 4.12 Comparison of telomere maps from normal and mutant alleles at the Xp/Yp telomere. (*A*) Telomere maps from tumour 955. The mutation found in the haplotype C-associated telomere consisted of 1 repeat insertion (underline). No mutations had been reported previously in the Xp/Yp haplotype C-associated allele using bulk DNA (Picket *et al.*, 2004). (*B*) Telomere maps from tumour 1314. Mutant allele 1 was found in 85% of the molecules screened at the haplotype B of the Xp/Yp telomere. This mutation has been described previously (Picket *et al.*, 2004). Mutant allele 2 shows a new mutation that arose on a copy of mutant allele 1 and it was present in 3/75 molecules analysed. Maps from the normal allele were determined using blood DNA from the same patient. In parenthesis at the end of each telomere map is shown the number of times the normal and the mutant alleles were observed. T, TTAGGG; G, TGAGGG; C, TCAGGG; N, repeats that do not amplify with the primers used.

The combined telomere mutation frequency of two tumours analysed in this chapter (tumours 955 and 1314) was statistically different compared to the mutation frequency of five tumours reported previously in Pickett *et al.*, 2004 (1.2% *vs* 35%; $P < 0.001$). The telomere mutation analysis carried out in Picket *et al.* was performed by comparing the telomere maps between the tumour DNA and blood DNA from the same patients whereas in this work the telomere mutation analysis was carried out by using single molecule STELA and TVR-PCR on tumour DNA. This indicates that the telomeres in the tumours analysed in this work continued to mutate but the mutation frequency decrease significantly once the tumours were established.

4.3.2.3 Telomere length in colon cancers

It has been reported previously that most colon tumours with MSI have shorter telomeres relative to the normal colon from the same patient. Furthermore, telomerase is often activated in these tumours suggesting telomere shortening may be a signal for telomerase activation in CRCs (Takagi *et al.*, 2000). Thus there could be a relationship between telomere shortening and CRC displaying MSI, suggesting a role of the mismatch repair system with telomere shortening.

Here, the telomere length from four pairs of matched normal colon and colon cancer samples were analysed by the STELA technique at the Xp/Yp telomere using the ImageQuant software as described above. All tumour samples showed MSI (assayed by instability in a tract of 10 adenines in the *TGFβRII* gene) and displayed telomerase activity (Pickett *et al.*, 2002). In all cases the telomere length was shorter in the tumours compared with its normal counterpart by an average of 1.5 kb (Figure 4.13 and Table 4.1). These results confirm previous reports where telomere shortening has been observed in such tumours (Hastie *et al.*, 1990; Takagi *et al.*, 2000). The telomere mutations found in two panels of CRC assayed by TVR-PCR on bulk DNA showed a tendency towards loss of telomere repeats (Pickett *et al.*, 2004) and thus it was suggested that an increase in telomere instability contribute to a more rapid telomere shortening in these tumours. Therefore it was hypothesised that the more rapid decline in telomere length in early stages of tumourigenesis in MMR defective tumours would increase the need to activate telomerase which in turn stabilises the telomeres and the genome.

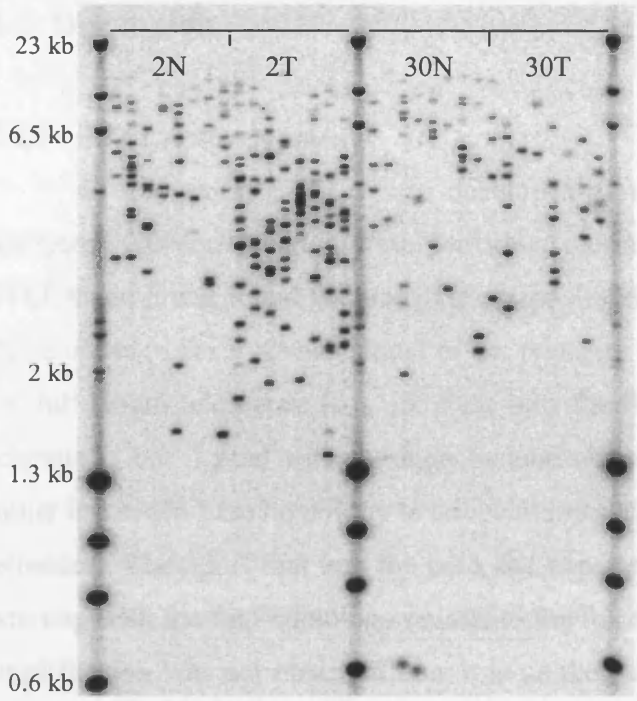


Figure 4.13 STELA at the Xp/Yp telomere from two pairs of normal and colon cancer samples. 250 pg of DNA were used per STELA reaction. Reduction in telomere length is observed in the colon cancer samples.

	2N	2T	23N	23T	26N	26T	30N	30T
Mean	5.37	3.49	4.26	3.21	5.13	3.7	5.59	4.1
Median	5.15	2.92	4.21	3.38	5.88	3.66	5.77	3.99
St Dev	2.51	2.27	1.56	1.69	2.16	1.93	2.86	2.1
St Error	0.34	0.32	0.18	0.17	0.56	0.26	0.45	0.38

Table 4.1 Telomere length of normal colon (N) and colorectal carcinoma (T) samples at the Xp/Yp telomere. All measurements are given in kb and where obtained using the ImageQuant software from Southern blots scanned with a phosphorImager.

4.4 Discussion

4.4.1 STELA development

In this work, STELA was developed at the polymorphic 16p/q, CB54 and 12q telomeres. Several controls were performed during the optimisation and development of STELA and it was found that one of the steps described in Baird *et al.* (2003), the ligation of the telorette to the 5' C-rich strand of the telomere, is not necessary for the amplification of the full length telomeres. It is not clear how the STELA products are generated when the telorette is not ligated with the high molecular weight DNA. However, since the telorette linker has seven-base homology to telomere repeats, it could itself act as a primer to amplify telomeres. Though if that was the case, the expected products would range from very short (starting with the first homology repeat) to the full telomere length. However, this pattern of amplification was not observed thus it is unlikely that telorette is acting as a primer in the reactions when the ligation step is not omitted.

A major step in the clarification of the importance of the ligation step came with the observation that hot-start PCR does not generate products if the telorette is not ligated to the C-strand of the telomere. This finding supports the idea of the need for the ligation step and so it still unclear how full length telomere amplification is achieved when the ligation step is omitted and hot-start is not used. Since the only difference with hot-start PCR is the addition of the DNA polymerases (*Taq* and *Pwo*) after the denaturation of the template, it is possible that one of these enzymes generates one strand of the full length telomere before the first denaturation step of the PCR (Figure 4.14). One possibility is that strand displacement occurs before the reaction is placed in the thermocycler, however strand displacement activity has not been reported for either of these enzymes. Another possibility is that *Taq* or *Pwo* polymerases ligate the telorette linker to the 5' end of the C-rich strand, though there are no reports to suggest that either *Taq* or *Pwo* have ligase activity (Figure 4.14).

4.4.2 Telomere stability in the colon

Using SM-STELA together with TVR-PCR it was possible to analyse the stability of telomeres in three normal colon samples. Although mutations at telomeres where found at two of them, they were identified in telomeres that contained blocks of the CTAGGG telomere-variant repeats and not in alleles without (discussed in Chapter 6). This suggests that most telomeres are fairly stable in the normal colon even though the samples were obtained from patients with more than 65 years of age. It has been found previously that

mutations at microsatellites accumulate with age in somatic tissue and although the mutation frequencies were low, mutations could only be detected using small-pool PCR techniques. For example, the mutation frequency in normal blood was assayed at mono- and di-nucleotide microsatellites and it was shown that in samples from 20-30 year-old individuals the mutation frequency was 0.9% per allele. However, the mutation frequency rose to 3.4% per allele in individuals that were 60-70 years old (Coolbaugh-Murphy *et al.*, 2005). Therefore, telomere mutations might also arise in the normal colon but the level of instability might be below what could be detected by screening about one hundred molecules with STELA and TVR-PCR. Further analysis would have to be carried out to establish whether instability at telomeres is present at a very low level in normal colon and if it is affected by the age of an individual.

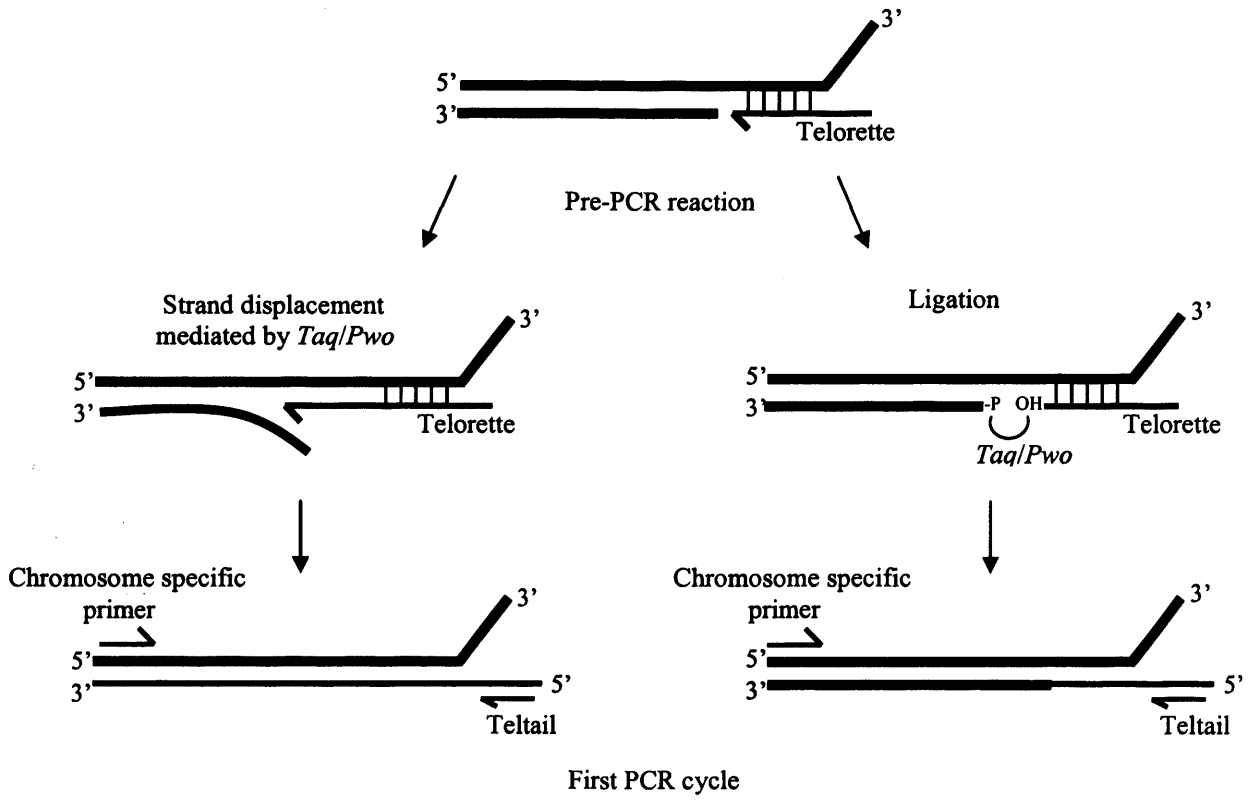


Figure 4.14 Models of telomere amplification using the STELA technique without hot-start PCR. Prior to the first cycle of amplification, *Taq* or *Pwo* polymerase might be involved in strand displacement of the C-strand or in the ligation of the telorette to the 5' end of the C-strand. This would allow amplification of the full length of a telomere.

4.4.3 Telomere instability in colon carcinomas

Telomeres from two colon carcinomas with characterised mutations at the *MSH2* gene were analysed by SM-STELA and TVR-PCR. The telomere mutation frequency of both tumours was similar (1.4% per allele). Previously the mutation frequency of 5 tumours with *MSH2* mutations (including both of the tumours analysed in this work) was calculated as 35% per allele (Pickett *et al.*, 2004). That study was performed by comparing the telomere maps from the tumours to blood DNA from the same patient. One explanation for the different mutation frequencies (35% and 1.4%) is that telomeres are highly unstable during tumour initiation but once the tumour is established the telomere mutation rate declines.

It has been proposed that genome instability is the main force that drives colorectal tumourigenesis (Nowak *et al.*, 2002; Calabrese *et al.*, 2004). In addition, most human cancers show chromosomal instability that generates imbalance in the number of chromosomes and contribute to loss of heterozygosity. In the colon it is believed that chromosomal instability is the first event during tumourigenesis leading to inactivation of tumour suppressor genes such as the *APC* (Nowak *et al.*, 2002). However, approximately 10% - 15% of sporadic CRCs do not present a high incidence of chromosomal instability but instead they show MSI caused by mutations in genes involved in the MMR system. Tumour progression in HNPCC patients suggests that there is an accumulation of mutations at oncogenes and/or tumour suppressor genes in stem cells at early stages of pre-tumour formation and that loss of MMR occurs late in life. However loss of MMR always occurs before clonal expansion (Calabrese *et al.*, 2004).

Thus, it is likely that the telomere mutations observed in the colon carcinomas originated at the time of tumour formation and most probably before clonal expansion, since most of the molecules analysed have the mutant telomere map. Only a small percentage (15% - 24%) of the telomere molecules analysed in two *MSH2* defective tumours retained the normal telomere map. This could be explained by the presence of normal cells in the tumour at the time the tumour was removed. Surprisingly, telomeres seem to be more stable once the tumour is established. Furthermore, a decrease in telomere length was observed between colon carcinoma and normal colon from the same patient. Thus it is possible that an elevated telomere mutation rate contributes to telomere shortening prior to telomerase activation. Furthermore, it was shown in this work that the mutation frequency of two *MSH2* deficient tumours (1.4% per telomere) and the mutation frequency in MMR deficient cell lines (3.6% per telomere) is statistically similar ($P = 0.27$). Those MMR defective cell lines were established from tumours and are telomerase positive thus supporting the idea that after

tumour formation and probably helped by the presence of the enzyme telomerase, telomere mutation frequency is low.

The results obtained in this work could be equivalent to previous reports of tumourigenesis in breast cancer (Chin *et al.*, 2004) where an increase in genomic instability is seen during the transition from benign to malignant growth. This transition was characterised by telomere shortening and an increase in genome instability observed by the presence of anaphase bridges. Telomere dysfunction was proposed to be the main reason for genome instability and most probably originated from telomere shortening; however once telomerase was activated, genomic stability was observed (Chin *et al.*, 2004).

Chapter 5

Downregulation of *MSH2* expression by shRNA

5.1 Introduction

5.1.1 RNA interference

Mammalian gene function has been studied mainly by disrupting murine genes, the use of antisense RNA, introduction of transgenes or by using knock-outs among other techniques. However, these methods take a long time and are expensive. A novel strategy to study gene function was developed from studies in plants and is called RNA interference (RNAi). RNAi in plants is mainly a defence mechanism used to protect their genomes from viral infection and from random integration of transposons (Baulcombe 1996; Tabara *et al.*, 1999, Gitlin and Andino 2003) and consists of the degradation of the exogenous RNA by a series of enzymes that are described below. The process of RNAi has been observed in other organisms such as *Caenorhabditis elegans* where it was found that mutation in some genes such as *MUT-2, MUT-7, RDE-2* and *RDE-3* increased the mobilization of transposons by reducing the RNAi pathway (Tabara *et al.*, 1999).

Downregulation of gene expression using RNAi was first used in *C. elegans* where it was shown that by injecting *C. elegans* cells with double stranded RNA homologous in sequence to the mRNA of the gene of interest, there was a reduction in the expression of that gene. However this process was not observed when only single-stranded RNA, sense or antisense RNA was injected (Fire *et al.*, 1998). The process of RNAi occurs endogenously and is mainly mediated by microRNAs and small interfering RNA (siRNA).

5.1.2 Processing of dsRNA produces microRNAs or siRNA

MicroRNAs (miRNAs) are 21-25 nucleotide small, non-coding RNAs that regulate gene expression at the post-transcriptional level. The first miRNA to be discovered was *LIN-4* in *C. elegans* (Lee *et al.*, 1993). *LIN-4* gene is transcribed into a 22 nt non-coding RNA that contains similarities with at least 7 regions of the 3'-untranslated region (3'UTR) of a gene involved in post-embryonic development called *LIN-14*. Once the *LIN-4* miRNA is bound to the 3' UTR of the *LIN-14* mRNA, it acts as a negative regulator of protein synthesis, preventing the translation of *LIN-14* and therefore controlling post-embryonic

development (Wightman *et al.*, 1991). miRNAs are processed in at least two steps, in the first one, nascent mi-RNA transcripts (primary-miRNA) are processed into ~70 nucleotides (pre-miRNA) catalysed by a RNase-III enzyme encoded by a gene called *DROSHA* (Figure 5.1). The pre-miRNAs are subsequently transported to the cytoplasm by Exportin 5 (Exp5) where they continue to be processed to miRNAs by a ribonuclease III (RNase III)-like enzyme called Dicer (Figure 5.1; Hammond *et al.*, 2000; Lee *et al.*, 2003; Lund *et al.*, 2004).

Dicer contains a catalytic RNAase III and a dsRNA-binding domain with the ability to cleave pre-miRNA into mature 21 - 25 nt miRNA (Lee *et al.*, 2002). In addition, Dicer can cleave long exogenous dsRNA into siRNA of 21 - 26 bp fragments with a 2-bp 3' overhang (Figure 5.1; Elbashir *et al.*, 2001). Dicer localises in the cytoplasm in mammalian cells, which suggests that RNAi is a process that occurs preferentially in the cytoplasm (Billy *et al.*, 2001). In some organisms including *D. melanogaster* and *Arabidopsis thaliana* more than one copy of the Dicer gene is found and it has been proposed that the different Dicer enzymes process different dsRNAs. For instance, in *Drosophila*, two Dicer enzymes (Dicer-1 and -2) are found and recognise different RNA substrates, miRNA and siRNA respectively (Lee *et al.*, 2004). Only one copy of the Dicer enzyme has been found in other organisms such as *C. elegans* and mammals.

5.1.3 Assembly of miRNA and siRNA into RNA silencing complexes

Once Dicer enzymes have generated either miRNA or siRNA from dsRNA, the cleaved products are loaded into a multiprotein RNA-induced silencing complex (RISC), whose main function is to facilitate the sequence-specific recognition of the target mRNA and/or its degradation (Hammond *et al.*, 2000). The RISC complex contains at least one member of the Argonaute (Ago) protein family whose role is to unwind the double-stranded miRNA or siRNA and to cleave the mRNA/siRNA complex (Song *et al.*, 2004). Mutation analysis of the carboxylase motif (Asp-Asp-Glu) of the human *AGO2* gene showed that it is required to cleave the targeted mRNA demonstrating that Ago2 protein is the main RISC component for the degradation of mRNA (Rivas *et al.*, 2005).

The number of Ago proteins differs between organisms, varying from one in *S. pombe*, 8 in humans and more than 20 in *C. elegans*. However, the function of most Ago proteins remains unknown (Sasaki *et al.*, 2003). In *D. melanogaster* and humans, Ago2 protein is the only known member of the Ago family that is required for siRNA-mRNA degradation. Therefore it is proposed that other members of the Ago family act in different steps of the RNA silencing pathway (Okamura *et al.*, 2004; Meister *et al.*, 2004).

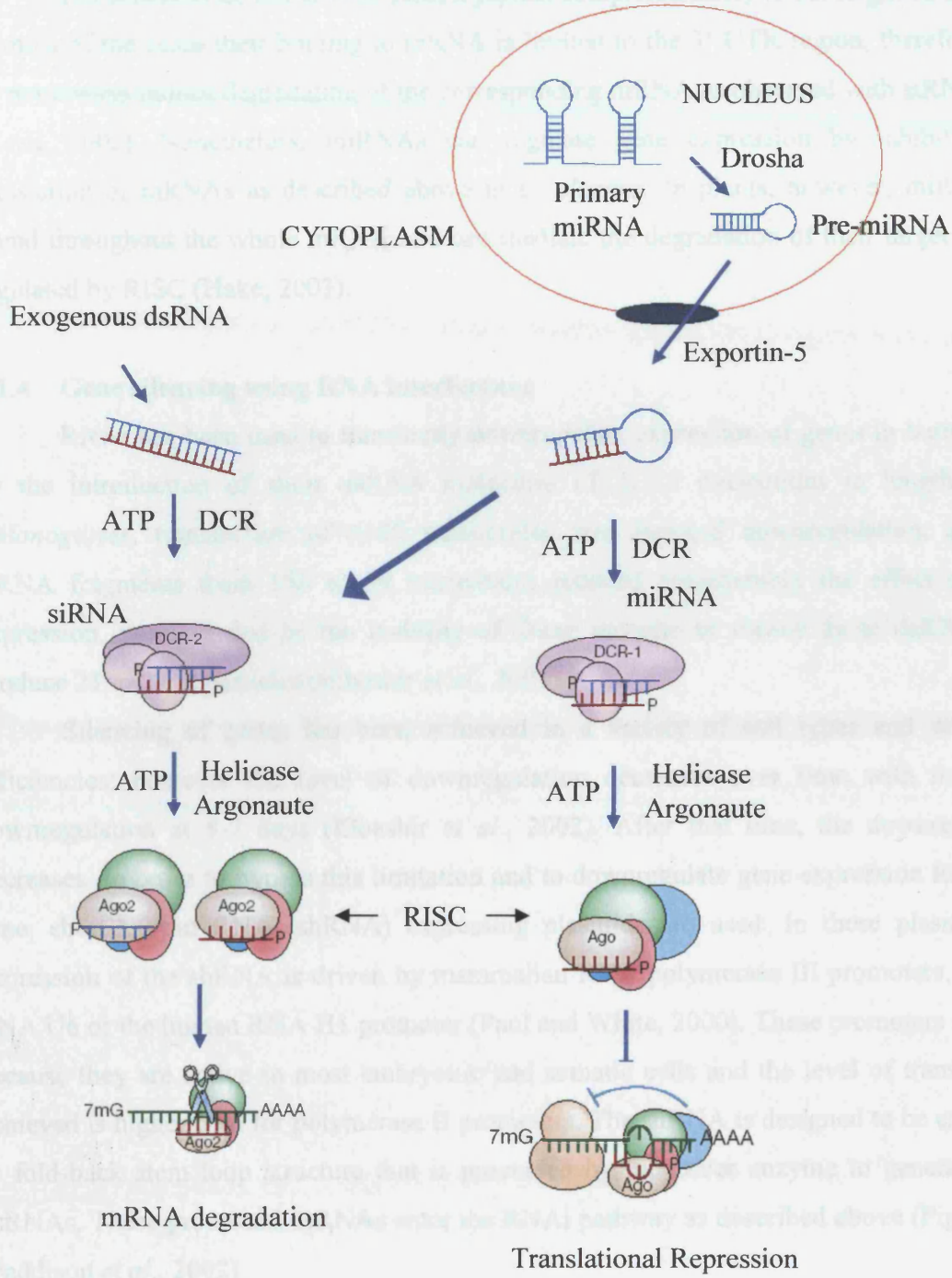


Figure 5.1 RNAi pathway. Genes that encode miRNAs are first transcribed into primary miRNA and processed in the nucleus by Drosha. The pre-miRNAs are exported to the cytoplasm where they are further processed by Dicer (DCR) to produce miRNAs. Dicer also processes exogenous dsRNA into siRNA. The duplex miRNAs or siRNAs are unwound when they are assembled into the miRNP or RISC complex. Argonaute proteins (Ago) mediate the translational repression (miRNA) or the target mRNA degradation (siRNA). Figure adapted from Meister and Tuschl, 2004.

The miRNAs do not always exhibit perfect complementarily to the target mRNA and in most of the cases their binding to mRNA is limited to the 3' UTR region, therefore they do not always induce degradation of the corresponding mRNA as observed with siRNA (Lee *et al.*, 2003). Nonetheless, miRNAs can regulate gene expression by inhibiting the translation of mRNAs as described above in *C. elegans*. In plants, however, miRNA are found throughout the whole mRNA and can mediate the degradation of their target mRNA regulated by RISC (Hake, 2003).

5.1.4 Gene silencing using RNA interference

RNAi has been used to transiently downregulate expression of genes in human cells by the introduction of short dsRNA molecules of 21-23 nucleotides in length. In *D. melanogaster*, transfection of >150 nucleotides also induced downregulation, although siRNA fragments from 150 to 39 nucleotides reduced considerably the effect on gene expression, possibly due to the inability of Dicer enzyme to cleave those dsRNAs and produce 21 to 22 nt particles (Elbashir *et al.*, 2001).

Silencing of genes has been achieved in a variety of cell types and with high efficiencies; however the level of downregulation decreases over time with maximum downregulation at 5-7 days (Elbashir *et al.*, 2002). After that time, the downregulation decreases. In order to bypass this limitation and to downregulate gene expression for longer time, short-hairpin RNA (shRNA) expressing plasmids are used. In these plasmids the expression of the shRNA is driven by mammalian RNA polymerase III promoters, such as RNA U6 or the human RNA H1 promoter (Paul and White, 2000). These promoters are used because they are active in most embryonic and somatic cells and the level of transcription achieved is higher than for polymerase II promoters. The shRNA is designed to be expressed as fold-back stem loop structure that is processed by the Dicer enzyme to generate short dsRNAs. These processed dsRNAs enter the RNAi pathway as described above (Figure 5.1) (Paddison *et al.*, 2002).

5.2 Aims

The Msh2 protein plays a major role in the MMR pathway; it initiates the recognition of base/loop mismatches and is involved in the recruitment of additional MMR proteins to initiate the repair process (Palombo *et al.*, 1996). Furthermore, it has been seen that tumours that harbour mutations in the *MSH2* gene show telomere instability detected by an increase in the mutation frequency (Pickett *et al.*, 2004).

Cell lines with defects in MMR genes established from tumours are used to study the roles of MMR genes in carcinogenesis. These cell lines show elevated rates of spontaneous mutations and genetic instability. These phenotypes could be a direct consequence of MMR defects however they could arise from the accumulation of alterations in other genes. In chapter 3 it was shown that the LoVo cell line (*MSH2*$^{-/-}$) has a telomere mutation frequency of 7.5% per haploid genome which is significantly higher than the mutation frequency in MMR proficient cell lines but lower than seen in tumours (Pickett *et al.*, 2004). Colon carcinomas and derived cell lines harbour numerous mutations and alterations in other genes that could mask the effects of MMR on telomere stability. In order to investigate this possibility, *MSH2* expression was downregulated using shRNA in a normal cell line and telomere stability and length were analysed.

5.3 Results

5.3.1 Immortalisation of CCD34-Lu cell line by *hTERT* transfection

The CCD34-Lu cell line was derived from the normal lung of a 2.5 week-old black child and it is a normal fibroblast cell line that senesces after approximately 40 population doublings. This cell line was selected for the downregulation of the *MSH2* gene because it contains telomeres that are suitable for telomere mutation detection by TVR-PCR. In order to have sufficient cells to perform the experiments required, the cell line was immortalised by transfecting it with a plasmid that contains a cDNA for the catalytic component of the telomerase, *hTERT*. The plasmid pCIneo-hTERT (Colgin *et al.*, 2000) (Appendix II) containing the *hTERT* cDNA was transferred into XL-1 blue *E. coli* competent cells as described in section 2.2.2. Colonies were selected in ampicillin, grown and plasmid mini-preps were used to recover pCIneo-hTERT plasmid. The plasmid was linearised using *Kpn*I restriction enzyme and purified by phenol/chloroform extraction prior to electroporation into CCD34-Lu cells.

The pCIneo-hTERT plasmid contains the neomycin phosphotransferase gene which confers G418 resistance on human cells that integrate the plasmid. Kill curves for CCD34-Lu cell line were therefore generated in order to determine the concentration of G418 required to kill non-transformed cells in approximately 7 days (Figure 5.2).

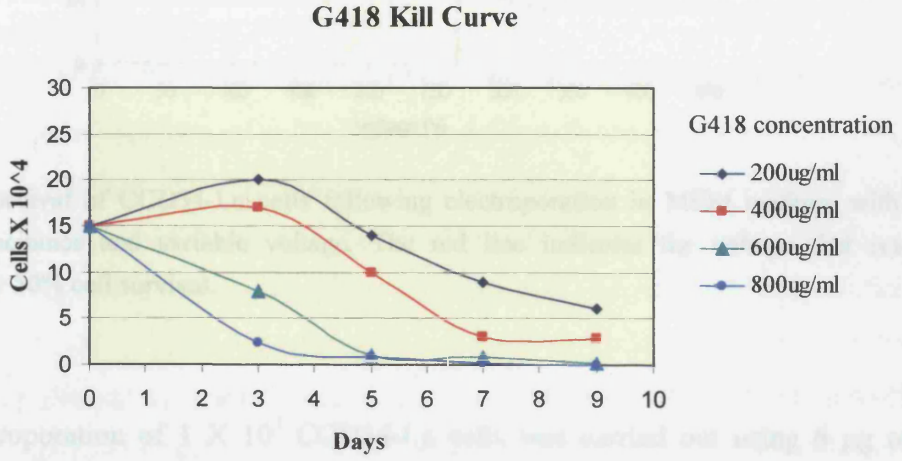


Figure 5.2 Survival of CCD34-Lu cells in medium containing G418. Kill curves were obtained using G418 with concentration ranging from 200 μg/ml to 800 μg/ml. The effective dose that kills non-transfected cells in approximately 7 days was 700 μg/ml.

Electroporation was used to deliver the pCIneo-hTERT plasmid into the CCD34-Lu cells. The optimum electroporation conditions vary between cell lines and it is reported that the conditions that produce approximately 50% cell survival provide an appropriate intake of exogenous DNA (Knutson and Yee, 1987; Chu *et al.*, 1987). Therefore, kill curves were established for CCD34-Lu using constant capacitance of 250 μF or 500 μF and using a high and low salt concentration medium (PBS and MEM). The use of PBS as the electroporation medium killed most of the cells in a narrow voltage window, therefore the conditions selected for electroporation of CCD34-Lu cell line in this work were 250 V and constant capacitance of 250 μF in MEM medium (Figure 5.3).

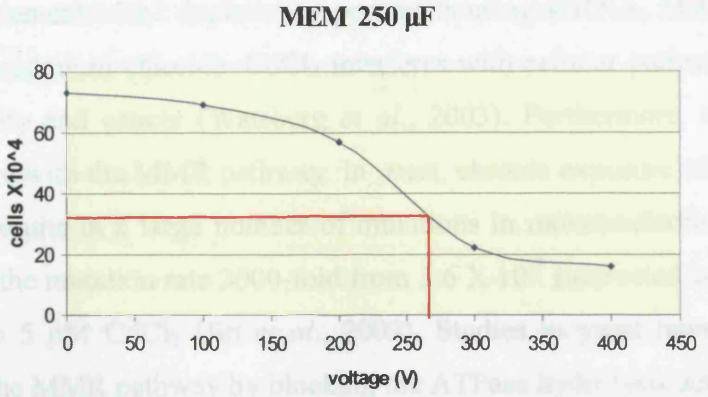


Figure 5.3 Survival of CCD34-Lu cells following electroporation in MEM medium with 250 µF constant capacitance and variable voltage. The red line indicates the voltage that resulted in approximately 50% cell survival.

Electroporation of 1 X 10^7 CCD34-Lu cells was carried out using 6 µg of linear plasmid (15 µl) under the conditions described in section 2.2.2.1 and the cells were plated in Petri dishes. The next day after electroporation, G418 was added to the medium and it was changed subsequently every 3 days until the end of the experiment. After 2 weeks, approximately 10 colonies were visible. The electroporation efficiency was calculated as 1.6 colonies/µg of DNA (or 1 colony per 1 X 10^6 cells). When the colonies reached a considerable size they were removed from the Petri dishes and were transferred to 6 well plates with 2 mls of medium, however most of the clones stopped dividing within a few days or never re-attached to the plate surface. The electroporation was carried out twice more using the same conditions as before and eventually two clones were recovered (CCD34-Lu/hTERT-1 and CCD34-Lu/hTERT-2). Another two electroporations were carried out with the conditions described above and they were plated in 80 cm^2 flasks. Following selection, 8-10 clones emerged in each flask and these were propagated as two mixed populations (CCD34-Lu/MixP1 and CCD34-Lu/MixP2). Telomerase expression was tested using the TRAP assay and both clones and the mix-populations were positive for telomerase activity, however all the experiments described below were performed with CCD34-Lu/hTERT-1. This clone was grown until it stopped dividing (40 to 50 PDs after *hTERT* integration) and it was assumed that *hTERT* was silenced during cell propagation.

5.3.2 Inhibition of Msh2 activity using cadmium chloride

To complement Msh2 depletion experiments using shRNA, Msh2 activity was also disrupted using cadmium chloride. $CdCl_2$ interferes with cellular pathways that can lead to genome instability and cancer (Waisberg *et al.*, 2003). Furthermore, this metal has been found to interfere with the MMR pathway. In yeast, chronic exposure of non-lethal doses of $CdCl_2$ (5 μM) results in a large number of mutations in mononucleotide runs in the *LYS2* gene increasing the mutation rate 2000-fold from 3.6×10^{-7} (untreated cells) to 7.8×10^{-4} in cells exposed to 5 μM $CdCl_2$ (Jin *et al.*, 2003). Studies in yeast have shown that $CdCl_2$ interferes with the MMR pathway by blocking the ATPase hydrolysis activities of the Msh2-Msh6 complex and therefore modulating the interactions of this complex with the mismatched DNA and other downstream factors (Banerjee and Flores-Rozas, 2005).

In this work CCD34-Lu/hTERT-1 cell line was exposed to various sub-lethal concentrations of $CdCl_2$. The cells were seeded at 2.5×10^5 cells per 25 cm^2 tissue culture flasks and were allowed to attach overnight. The next day, new media containing 0.5 to 15 μM $CdCl_2$ was added and the cells allowed to grow to confluence. Each time the cells were trypsinised and replated $CdCl_2$ was added to the medium. Cells exposed to $CdCl_2$ with concentrations of 10 μM or higher started to die just a few population doublings after the addition of $CdCl_2$, therefore these cells were not used for subsequent MSI and telomere mutation analysis.

It has been shown that defects in MMR genes lead to MSI (Aaltonen *et al.*, 1993). As a consequence, MSI can be used as a marker to detect deficiencies in the MMR pathway. In order to determine whether $CdCl_2$ affected the MMR system, MSI analysis was carried out in cells exposed to concentrations of 0.5, 1 and 5 μM by testing the stability of Bat-26. The BaT-26 microsatellite consists in a poly(A) tract located in intron five of the *MSH2* gene. This microsatellite is monomorphic and it is highly stable at both alleles within an individual and between individuals, however it is unstable in the majority of tumours and cell lines with defects in MMR. Therefore it is possible to use Bat-26 alone as a marker for MSI (Hoang *et al.*, 1997; Bocker *et al.*, 1997). It has been reported that a mutant allele must be present at a frequency >0.25 to be detected by PCR in bulk genomic DNA (Coolbaugh-Murphy *et al.*, 2004). Therefore, Bat-26 stability was assayed by small-pool PCR to allow detection of low-frequency mutant alleles as described in section 2.2.7.1.

MSI was not detected in samples exposed to 0.5 μM and 1 μM doses; however it was found in cells exposed to 5 μM $CdCl_2$ after 18 PDs. The mutation frequency at the Bat-26 microsatellite was determined as 4.4% (4/90 or 2.4×10^{-3} per PD), which was significantly

higher than in control cells without any treatment (0/120), $P = 0.035$. These results show that $CdCl_2$ can induce low levels of MSI in CCD34-Lu cell line exposed to sub-lethal concentrations of cadmium and that the mutations could be detected using a small-pool PCR approach.

To determine whether $CdCl_2$ induce instability at telomeres, TVR-PCR from single molecule STELA products were performed in the cells that showed MSI (cells exposed to 5 μM $CdCl_2$ for 18 PDs). Before mutation analysis was performed, the haplotype of the Xp/Yp and 12q-adjacent telomeres was determined and telomere maps were obtained by TVR-PCR using genomic DNA from CCD34-Lu cell line to identify telomeres with interspersion patter of repeats that would facilitate the detection of mutations. Haplotype analysis at the Xp/Yp-adjacent telomere showed that the CCD34-Lu cell line was heterozygous for haplotype A and D. Analysis at the 12q telomere-adjacent sequence suggests the CCD34-Lu cell line is homozygous (B/B alleles). This cell line was also tested for the presence of the polymorphic end at the 16p/q telomere but it does not contain that sequence. TVR-PCR analysis revealed that Xp/Yp and 12q telomeres were suitable for mutation analysis.

In total 76 alleles were screened for mutations at the Xp/Yp telomere (haplotype A) and 73 alleles (haplotype B) at the 12q telomere, however no mutations were detected. These experiments demonstrate that low doses of $CdCl_2$ can induce MSI in CCD34-Lu/hTERT-1 cell line, though the effect of 5 μM exposure is not enough to destabilise telomeres after 18 PDs. It is possible that longer exposure times would increase MSI and telomere mutations. However in this work it was not possible to keep the cells dividing for longer since they stopped dividing at around PD 20 after $CdCl_2$ was added.

5.3.3 Transient downregulation of *MSH2* using siRNA

In order to achieve long term downregulation of *MSH2* expression, shRNA was used. However, it is reported that variation in the level of downregulation depends on the sequence of the siRNA. Therefore before any shRNA experiments were carried out, transient downregulation by siRNA was preformed. Two duplex siRNA were used for transient *MSH2* downregulation. The first sequence was designed using the Oligoengine RNAi design software (siRNA 1; sequence found at exon 6 of the *MSH2* gene) while another sequence (siRNA 2; sequence found at exon 2 of the *MSH2* gene) was already described in the literature (Wang and Qin, 2003). A blast-search was performed with both siRNA in order to determine their specificity and to avoid targeting other genes. Both sequences are formed by 21 nucleotide sense and antisense strands containing a 2 nucleotide thymidine 3′ overhang to

give stability to the RNAs (Figure 5.4). The siRNA sequences were commercially synthesised and annealed (Ambion). Electroporation was used to deliver the annealed siRNA into a control cell line HT1080 (telomerase positive cell line). Two different amounts of dsRNAs were used (300 or 600 pmols), the cells were collected after 46 and 72 hours of electroporation and total protein were extracted. Western blotting was carried out to measure Msh2 protein expression in the presence of siRNA 1 or 2 after 46 and 72 hrs. A significant decrease in protein expression was observed after 46 and 72 hrs using 600 pmols of siRNA 1 but not for siRNA 2 (Figure 5.4). Thus the sequence from siRNA 1 was used to design a shRNA.

A)

```
3'-   GCCUCUCAUGGAUAAGAACTT - 5'    siRNA 1
5'- TTCGGAGAGUACCUAUUCUUG   - 3'

3'-   AAUCUGCAGAGUGUUGUGCTT - 5'    siRNA 2
5'- TTUUAGACGUCUCACAACACG   - 3'
```

B)

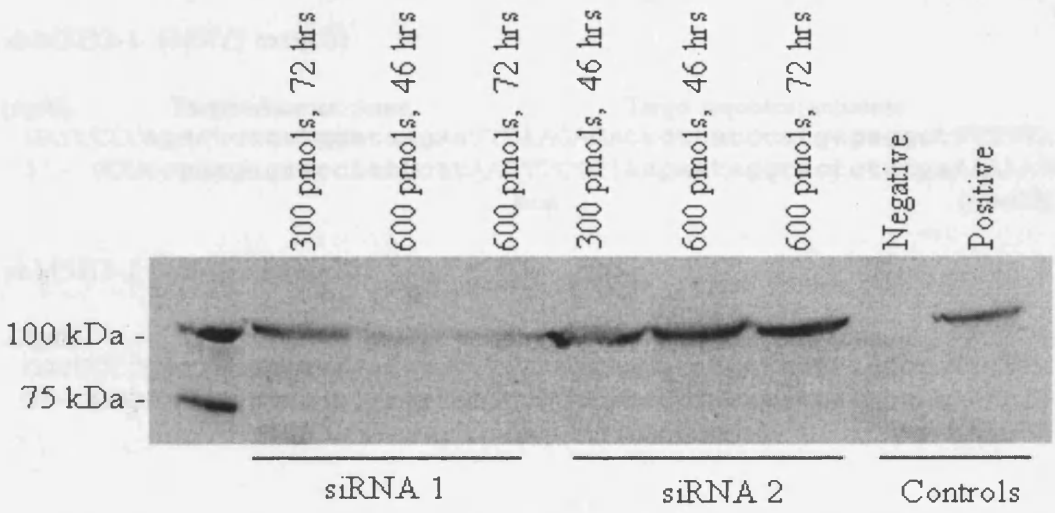


Figure 5.4 Transient RNAi assays were used to test the effectiveness of siRNA1 and siRNA2 at depleting the level of Msh2 protein. A) Oligonucleotide sequence used for siRNA experiments. B) Western blotting of 30 μg of protein extract incubated with a mouse antibody against Msh2. The amount of siRNA used together with the time after transfection are given. LoVo (*MSH2*$^{-/-}$) cell line was used as a negative control and HT1080 (without any treatment) as a positive control. All experiments were carried out using HT1080 cell line.

5.3.3.1 shRNA design and vector preparation

pSuperior.puro vector (Oligoengine, map is shown in Appendix II) was chosen for delivery and expression of the shRNAs. This vector is 4.3 kb in length and contains a polymerase-III RNA promoter (H1 promoter) and the puromycin resistant gene for selection. Two DNA oligonucleotides, for targeting *MSH2* by shRNA, were designed to be inserted into pSuperior.puro. The first shRNA plasmid was based on sequence 1 from the siRNA experiments and the second sequence was generated using the RNAi codex design software (Figure 5.5) (Olson *et al.*, 2001). Both shRNA oligos contained a 19-nucleotide sense and antisense sequence derived from *MSH2* mRNA, separated by a 9-nucleotide loop. The oligos also contained *Bgl*II and *Hind*III restriction site-overhangs to be used for cloning. The sense and antisense oligonucleotides were diluted to 3 μg/ml and annealed. The prepared double stranded oligos were ligated to approximately 5 μg of plasmid (linearised with *Bgl*II and *Hind*III restriction enzymes) and subsequently purified by phenol/chloroform extraction (Figure 5.6).

shMSH2-1 (*MSH2* exon 6)

(*Bgl*II) Target sequence: sense Target sequence: antisense

*GATC*CCC**agcctctcatggataagaa**TTCAAGAGA**ttcttatccatgagaggct**TTTTTA – 3′

3′– GGG**tcggagagtacctattctt**AAGTTCTCT**agaataggtactctccga**AAAAAT*TCGA*

loop (*Hind*III) →

shMSH2-2 (*MSH2* exon 13)

(*Bgl*II) Target sequence: sense Target sequence: antisense

*GATC*CCC**gaggtaaatcaacatatat**TTCAAGAGA**atatatgttgatttacctc**TTTTTA – 3′

3′– GGG**ctccatttagttgtatata**AAGTTCTCT**tatatacaactaaatggag**AAAAAT*TCGA*

loop (*Hind*III) →

Figure 5.5 Double-stranded oligonucleotides used to deplete *MSH2* expression. In bold are the sense and antisense RNAi sequences. shMSH2-1 is complementary to a sequence in exon 6 while shMSH2-2 to exon 13 of *MSH2*. *Bgl*II and *Hind*III restriction sites are indicated. The arrow shows the direction of the transcription.

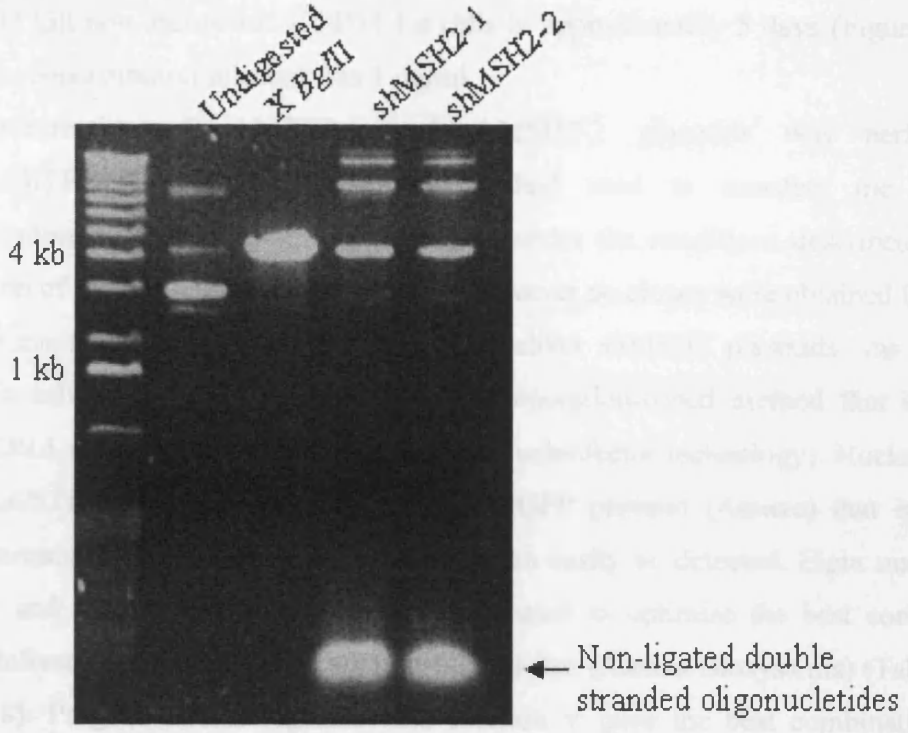


Figure 5.6 The pSuperior.puro plasmid ligated to the shRNA DNA oligonucleotides shMSH2-1 and shMSH2-2. The pSuperior.puro plasmid is shown undigested, digested with *Bgl*II (linear plasmid) and following ligation to shMSH2-1 and shMSH2-2.

Successful cloning of the oligonucleotides into pSuperior.puro should destroy the *Bgl*II restriction site; therefore prior to transformation into *E. coli*, the plasmids were digested with *Bgl*II in order to reduce false positive colonies. Each plasmid (30 ng) was electroporated into the XL1 blue *E. coli* and selected on ampicillin plates (Sambrook *et al.*, 2001). Colony PCR was used to screen for clones containing the oligo inserts using primers pSUP-1 together with pSUP-2. Positive colonies were also confirmed by sequencing using the PRSuper primer (Appendix III). shMSH2-1 and shMSH2-2 plasmids with the right insert were prepared for transfection into human cells by linearization with *BamH*I and purified by phenol/chloroform extraction. The DNA was ethanol precipitated and the pellet resuspended in sterile water.

5.3.3.2 CCD34-Lu/hTERT cells transfection with shMSH2 plasmids

pSuperior.puro plasmid contains the puromycin resistant gene as a selectable marker for mammalian cells, therefore a kill curve was established to determine the antibiotic dose

that would kill non-transfected CCD34-Lu cells in approximately 5 days (Figure 5.7). The puromycin concentration selected was 1 μg/ml.

Transfection of shMSH2-1 and shMSH2-2 plasmids was performed in CCD34Lu/hTERT-1 cell line. The first method used to transfect the cells was electroporation. Electroporation was carried out under the conditions described above for transfection of the *hTERT*-containing plasmid; however no clones were obtained from any of the experiments. Therefore another method to deliver shMSH2 plasmids was used. This method is called nucleofection and is an electroporation-based method that can deliver plasmid DNA directly into the nucleus (Amaxa Nucleofector technology). Nucleofection of CCD34-Lu/hTERT-1 was optimised using pmaxGFP plasmid (Amaxa) that encodes the green fluorescent protein, thus transfected cells can easily be detected. Eight nucleofection programs and two nucleofection solutions were used to optimise the best conditions for plasmid delivery into the CCD34-Lu/hTERT-1 cell line (Amaxa Biosystems) (Table 5.1 and Figure 5.8). Program X-005 together with solution V gave the best combination of cell survival and plasmid uptake.

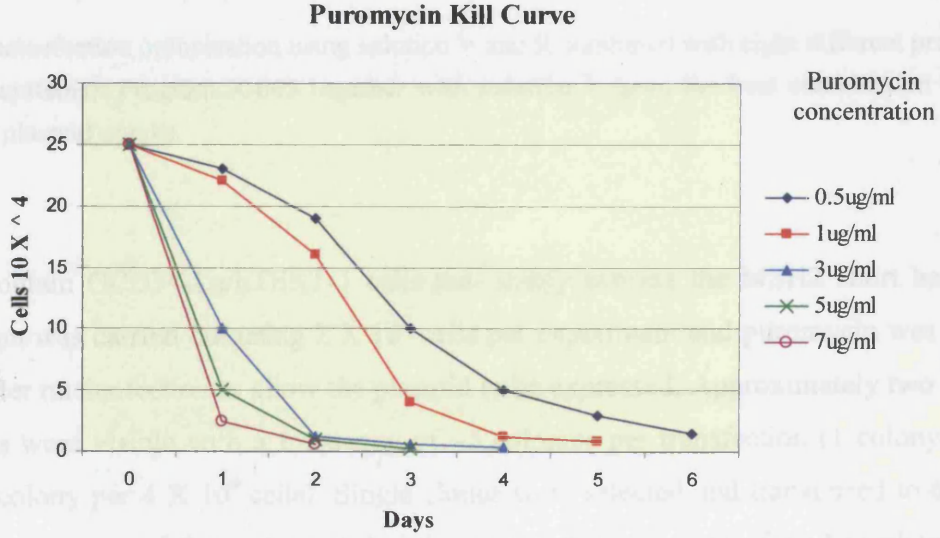


Figure 5.7 Kill curve of CCD34-Lu cells in medium containing puromycin. Concentrations ranging from 0.5 μg/ml to 7μg/ml were used. It was found that 1 μg/ml killed non-transfected cells in approximately 5 days.

Solution V

Programme	Total screened cells	% of transfected cells	% surviving cells
U-23	284	37	24
U-012	235	29	13
T-020	236	35	11
T-030	265	45	10
X-001	214	30	23
X-005	**216**	**39**	**30**
L-029	290	39	29
D-023	220	33	23

Solution R

Programme	Total screened cells	% of transfected cells	% surviving cells
U-23	262	36	24
U-012	244	30	27
T-020	230	33	20
T-030	225	26	19
X-001	258	24	42
X-005	218	29	21
L-029	215	20	13
D-023	248	21	24

Table 5.1 Nucleofection optimisation using solution V and R combined with eight different programs (Amaxa Biosystems). Program X-005 together with solution V gave the best combination of cell survival and plasmid uptake.

To obtain CCD34-Lu/hTERT-1 cells that stably express the MSH2 short hairpins, nucleofection was carried out using 2 X 10^5 cells per experiment and puromycin was added two days after nucleofection to allow the plasmid to be expressed. Approximately two weeks later, clones were visible with a frequency of ~5 colonies per transfection (1 colony/μg of DNA or 1 colony per 4 X 10^4 cells). Single clones were selected and transferred to 6 well-plates; however most of them stopped dividing soon after being transferred to plates. The experiments were repeated several times and eventually five clones containing the shMSH2-1 plasmid were obtained. In order to confirm the integration of plasmid, genomic DNA was extracted from each clone and PCR was performed with primer pSUP-1 and pSUP-2 that amplify the shRNA insert (Figure 5.9).

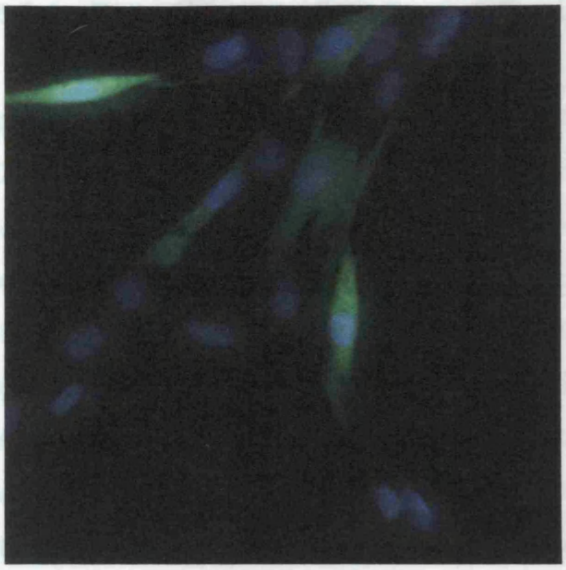


Figure 5.8 Detection of plasmid uptake after transfection of CCD34-Lu/hTERT-1 with the pmaxGFP plasmid. Cells that express the GFP protein (green) can be counted. In this experiments the pmaxGFP plasmid was transfected into the CCD34-Lu/hTERT-1 cells using solution V and programme X-005. The cells were fixed, the nuclei were stained with DAPI and the cells visualised using a fluorescent microscope.

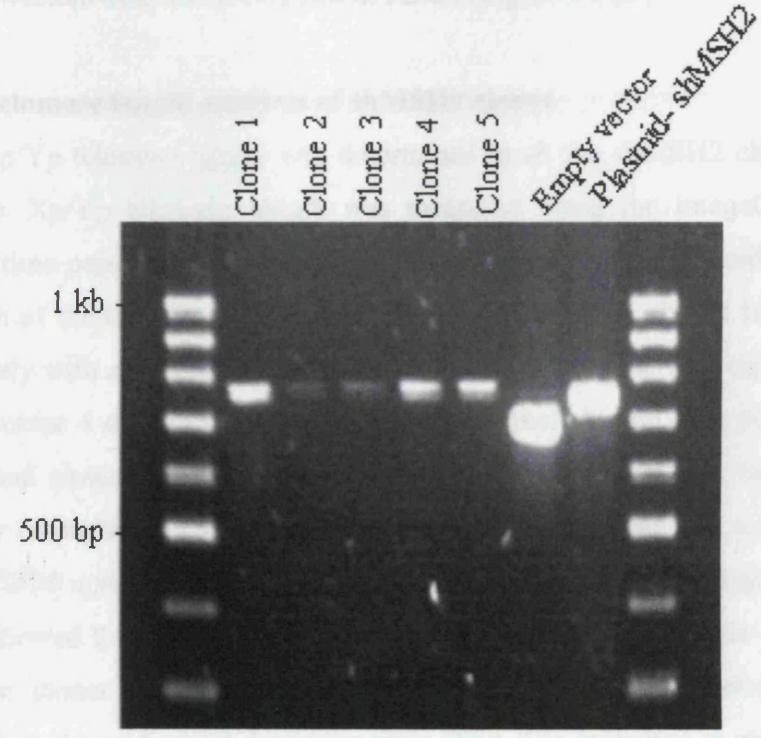


Figure 5.9 PCR of genomic DNA from CCD34-Lu/hTERT-1 clones that contain the shMSH2-1 plasmid. An empty vector (pSuperior.puro) was used as a negative control and DNA from the shMSH2-1 plasmid was used as a positive control. PCR primers pSUP-1 and pSUP-2 were used to amplify the shRNA insert.

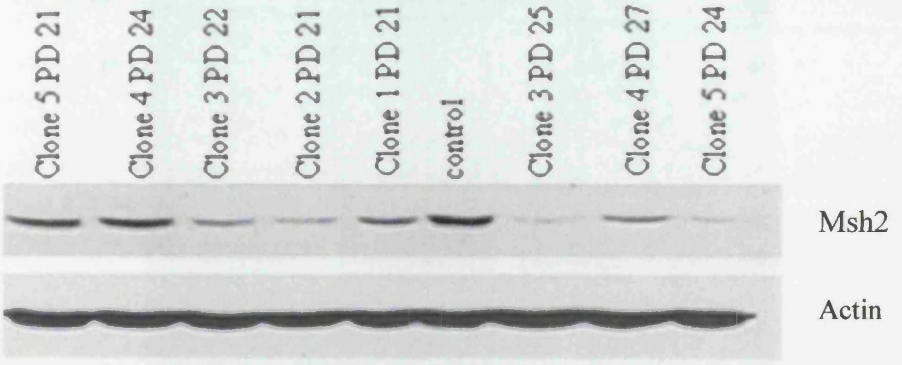
5.3.4 Detection of Msh2 downregulation in shRNA positive clones

Cells from the 5 different shMSH2-1 containing clones were harvested and total proteins extracted. For clones 1 and 2, cells were harvested once they reached confluence in 25cm^2 flasks, representing approximately 21 PDs. For clones 3, 4 and 5, cells were harvested at two points during propagation. Western blotting was performed on all clones and in most cases the expression of *MSH2* was determined from two independent experiments using ImageQuant software and normalised using the expression of actin as a loading control (Figure 5.10 A). The signal of CCD34-Lu/hTERT-1 cell line without any short hairpin was set as 100% and from that value the percentage of *MSH2* expression was calculated for all clones (Figure 5.10 B). The highest downregulation was observed in clones 2 and 3, therefore these two clones were used to assess MSI and telomere stability (section 5.3.7). Even though clone 5 (PD 24) showed a downregulation of ~ 90 %, this clone was excluded from telomere mutation analysis since at PD 21 it showed a downregulation of ~40% and only one western blot was performed at PD 24 (Figure 5.10B).

5.3.5 Telomere length analysis of shMSH2 clones

Xp/Yp telomere length was determined in all the shMSH2 clones using the STELA technique. Xp/Yp telomere length was measured using the ImageQuant software at two different time points (3 PDs apart) in clones 3, 4 and 5. It was found that in clones 3 and 5 the length of the telomeres decreased from 6.15 kb to 5.12 kb and from 4.79 kb to 3.38 kb respectively with an average of ~350 bp and ~450 bp loss per PD respectively (Figure 5.11), however clone 4 did not show any change in telomere length. The mean telomere length of clone 3 and clone 5 was significantly different compared to their mean telomere length 3 PDs later respectively (Student's *t* test, $P > 0.001$) whereas the mean telomere length of clone 4 PD 26 and PD 29 was similar (Student's *t* test, $P = 0.921$; Figure 5.11). Interestingly, clone 4 showed the highest Msh2 expression; therefore it is possible that depletion of Msh2 protein in clones 3 and 5 affected Xp/Yp telomere length. Telomere length was also measured in clones 1 and 2, however since DNA was extracted at one time point it was not possible to determine whether telomere length was eroded during replication in these clones.

A)



B)

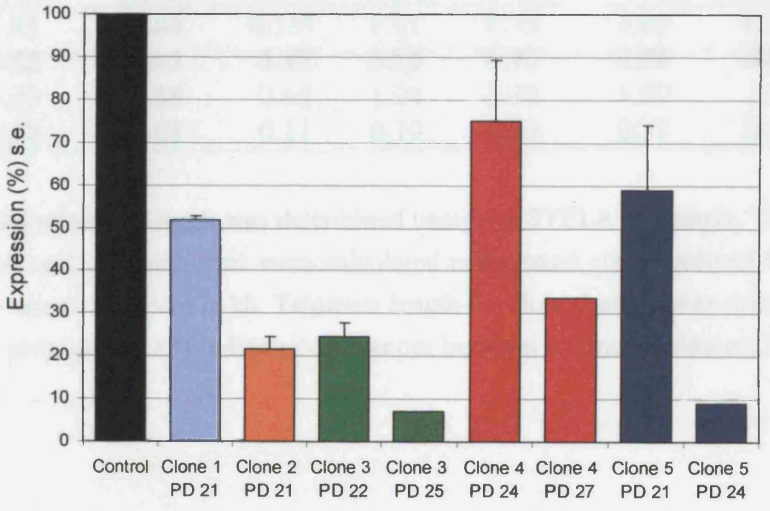
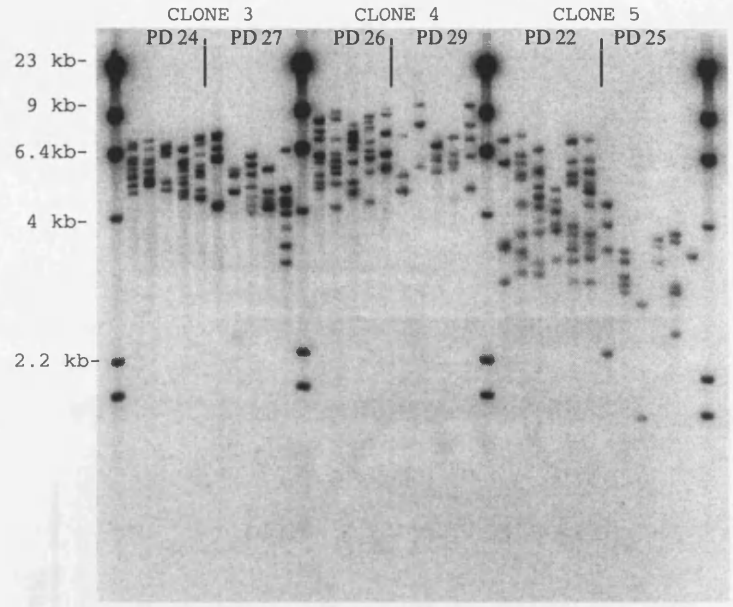


Figure 5.10 Analysis of *MSH2* expression after shMSH2-1 transfection. (*A*) Msh2 western blotting of 5 clones transfected with shRNA. Protein expression was quantified using actin as a loading control. (*B*) *MSH2* expression is presented as a percentage of the level of expression in the control (set to 100%). Standard error (s.e.) bars are shown for samples were more than one western blot was performed.



	Clone 1	Clone 2	Clone 3		Clone 4		Clone 5	
	PD 22	PD 21	PD 24	PD 27	PD 26	PD 29	PD 22	PD 25
Mean	5.83	4.64	6.15*	5.12	6.76	6.80	4.79*	3.38
Median	5.82	4.65	6.12	5.29	6.47	6.46	4.65	3.49
St Dev	1.06	0.58	0.65	1.04	1.38	1.60	1.28	0.67
St Error	0.17	0.08	0.11	0.19	0.23	0.37	0.18	0.15

Figure 5.11 Xp/Yp telomere length was determined using the STELA technique. The mean, median, standard deviation and standard error were calculated using band sizes obtained from ImageQuant software. All the values are given in kb. Telomere length for clone 1 and 2 was calculated at only one point during cell propagation. (*) indicates differences between the mean telomere length 3 PDs later (Student's *t* test).

It is possible that the variation in telomere length between clones is due to clonal variation, though all clones derived from a single clone (CCD34-Lu/hTERT-1). Another possibility is that these clones have different levels of telomerase activity. Therefore the Telomerase Repeat Amplification Protocol (TRAP; Kim *et al.*, 1994) assay was performed using a kit (TRAPeze telomerase detection kit, Intergen) in shMSH2-1 clones 3, 4 and 5 to check telomerase activity. Quantitative analysis of telomerase expression in clones 3, 4 and 5 showed similar level of expression in all clones with 98, 106 and 88 units of total product generated (TPG; 1 unit correspond to the product generated by approximately 33 telomerase positive control cells) (Figure 5.12). The determination of the TPG units was performed using the following equation:

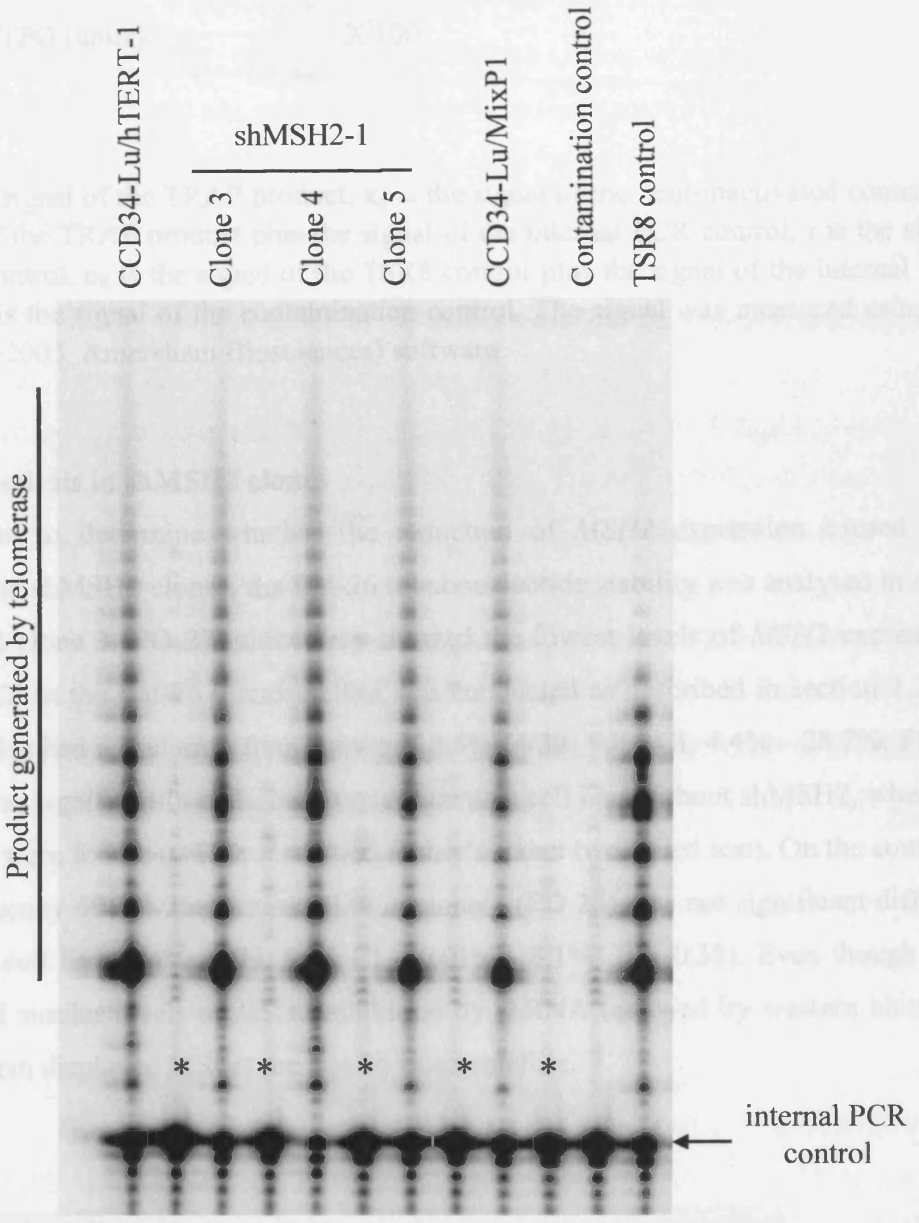


Figure 5.12 Telomerase activity assayed using the TRAP assay. All shMSH2-1 clones were generated from CCD34-Lu/hTERT-1. For each sample a heat inactivation control was performed (*). In the contamination control the cell extract was substituted by lysis buffer. TSR8 is an oligonucletide with 8 telomeric repeats and was used as a control template instead of sample extract.

$$\text{TPG (units)} = \frac{(x - x_o)/\, c}{(r - r_o)/\, c_R} \text{ X } 100$$

where x is the signal of the TRAP product, x_o is the signal of the heat-inactivated controls, c is the signal of the TRAP product plus the signal of the internal PCR control, r is the signal of the TSR8 control, c_R is the signal of the TSR8 control plus the signal of the internal PCR control and r_o is the signal of the contamination control. The signal was measured using the ImageQuant (v2005, Amersham Biosciences) software.

5.3.6 MSI analysis in shMSH2 clones

In order to determine whether the reduction of *MSH2* expression caused MSI instability in the shMSH2 clones, the Bat-26 mononucleotide stability was analysed in clone 2 (PD 21) and clone 3 (PD 27) since they showed the lowest levels of *MSH2* expression. Small-pool PCR at the Bat-26 microsatellite was conducted as described in section 2.2.7.1. Clone 2 (PD 21) had a mutation frequency of 12.5% (4/32; 95% CI, 4.4% - 28.7%; Figure 5.13) which was significantly different from the control cell line without shMSH2, where no mutant alleles were found, (0/120; $P = 0.005$ Fisher's Exact two-tailed test). On the contrary, mutation frequency of Bat-26 microsatellite in clone 3 (PD 25) was not significant different to the control cell line, 1.4% (1/73, 95% CI, <0.01% - 8.1%; $P = 0.38$). Even though both clones showed similar levels of *MSH2* inhibition by shRNA (assayed by western blotting), only one of them displayed MSI at the Bat-26 microsatellite.

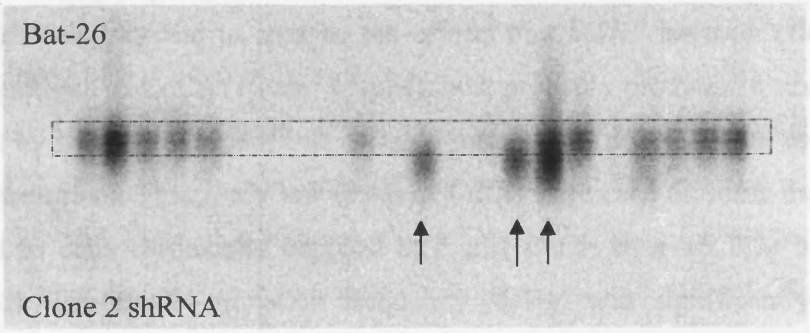


Figure 5.13 Small-pool PCR at the Bat-26 microsatellite. Dilutions of DNA were carried out to get approximately one amplifiable molecule of DNA per PCR reaction. Mutant alleles found in clone 2 are shown with arrows.

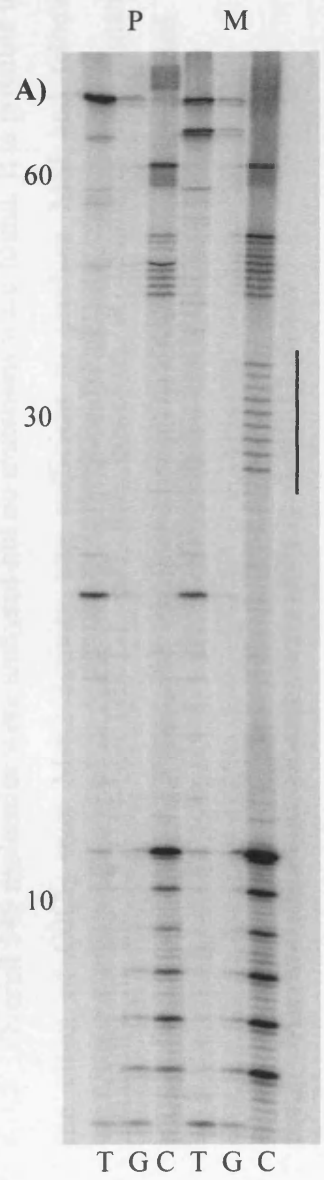
5.3.7 Telomere mutation detection in shRNA clones

As mention above, telomere maps of CCD34-Lu cell line at the Xp/Yp and 12q telomeres consisted of an interspersion of TTAGGG and variant repeats. The haplotype D-associated telomere was used for mutation detection analysis in shMSH2 clones 2 and 3. Single molecule STELA was carried out using allele specific primers to amplify this telomere as described previously. In total, 76 molecules were analysed from clone 2 (PD 21) and 1 telomere mutation was found. The mutation frequency was 1.3% per allele (95% CI, 0.3% - 7.1%) or 2.5% per kb per allele. The telomere mutation frequency found in clone 2 was similar to that found in the *MSH2*$^{-/-}$ LoVo cell line (P = 0.14; Fisher's exact two-tailed test). The mutation found comprises the appearance of a block of 9 TCAGGG repeats in a place where no telomere variant repeats amplify with the repeat primers used (N-type repeats). This mutation did not affect the surrounding repeats therefore there is no gain or loss of repeats. However, the progenitor allele (and mutant) contains another block of 9 TCAGGG repeats six repeats further into the telomere repeat array. Thus, it is possible that this mutant originated by an intra-allelic gene conversion-like or a unequal sister chromatid exchange process (Figure 5.14). Similar analysis was carried out in clone 3 (PD 27), however, no mutations were found in 80 molecules screened at the Xp/Yp haplotype D-associated telomere (0/80).

5.4 Discussion

5.4.1 The use of $CdCl_2$ to inhibit Msh2 activity

$CdCl_2$ is a mutagen that interferes with Msh2 activity and affects the recruitment of downstream MMR factors (Banerjee and Flores-Rozas, 2005). In this work $CdCl_2$ was used to inhibit Msh2 activity and to analyse the effects that *MSH2* have on telomere stability. High concentrations of $CdCl_2$ (10 and 15 μM) most probably produced a toxic effect on the cells, as seen by the slower rate of cell division soon after $CdCl_2$ addition compared to cells without any treatment. Thus, only low doses of $CdCl_2$ were used on these experiments. MSI was detected in cells chronically exposed to 5 μM $CdCl_2$ over 18 PDs but not at lower concentrations. The Bat-26 mutation frequency (4.4%) was significantly different to a control cell line. These findings indicate that sub-lethal concentrations of $CdCl_2$ probably alter Msh2 activity and induce MSI in CCD34-Lu/hTERT cell line. Telomere stability was analysed at the Xp/Yp and 12q by the used of TVR-PCR of single telomere molecules.



B)

```
          10                  30                  50                  70
          .                   .                   .                   .
P  GTTTTCCCCCCNNNNNNNNNTNTNNNNNNNNNNNNNNNTNNNCCCCCCCCCCNNNNTNNTCCCNNCNNNNNNNTTNNNNNNNNNTTNN..
M  GTTTTCCCCCCNNNNNNNNNTNTNNNNCCCCCCCCCCNNTNNNCCCCCCCCCCNNNNTNNTCCCNNCNNNNNNNTTNNNNNNNNNTTNN..
```

Figure 5.14 Telomere mutation at the Xp/Yp telomere (haplotype D) of clone 2 shMSH2-1. (A) TVR-PCR and (B) telomere maps are shown for the progenitor (P) and the mutant molecule (M). The mutation is either indicated by a bar in *A* or underlined in *B*. T, TTAGGG; G, TGAGGG; C, TCAGGG; N, non-amplifiable repeats or difficult to determine.

Overall 149 molecules were analysed but no mutations were found. It is possible that the effect of $CdCl_2$ over Msh2 activity was sufficient to produce MSI and telomere instability in the CCD34-Lu/hTERT-1 cell line however, at telomeres the mutation frequency might be lower and cannot be detected with the number of molecules screened.

5.4.2 Microsatellite and telomere instability in shMSH2 clones

In 1997, the National Cancer Institute (US) defined uniform criteria for the determination of MSI (Boland *et al.*, 1998) and among several parameters it was recommended the use of a reference panel of microsatellites. These microsatellites include two mononucleotides (Bat-26 and Bat-25) and three dinucleotides (D5S346, D2S123 and D17S250). In this work only Bat-26 microsatellite was used for the determination of MSI as it is unstable in nearly all samples that show MSI.

MSI and telomere instability were assayed in two clones expressing the shMSH2-1 short-hairpin (clone 2 PD 21 and clone 3 PD 27). MSI (at the Bat-26 microsatellite) and telomere instability (at the Xp/Yp telomere) was detected in clone 2 but not in clone 3 although both clones showed levels of *MSH2* downregulation $\geq$ 80% (Figure 5.10). The lack of MSI in clone 3 should be interpreted cautiously as only one microsatellite was assessed. Furthermore it is possible that MSI occurred in clone 3 but at a rate lower than could have been detected by small-pool PCR analysis used on Bat-26 (95% CI, <0.01% - 8.1%). Although *MSH2* downregulation assessed at ~PD 22 was similar between clones 2 and 3, it is not known what the downregulation status was at earlier stages of clonal expansion. Mutations at microsatellites are thought to accumulate and evolved in a stepwise manner and in poly(A) tracts such as Bat-26 there is a bias to deletion of repeats over time (Blake *et al.*, 2001). Thus, if the level of *MSH2* downregulation is maintained, it is predicted that MSI at Bat-26 would emerge in clone 3 after further rounds of replication.

In summary this work has shown that downregulation of *MSH2* expression in a clone of CCD34-Lu/hTERT normal fibroblast cell line induced MSI and telomere instability. Furthermore, the level of telomere instability correlates with the level seen in the *MSH2*$^{-/-}$ LoVo cell line ($P = 0.14$). This suggests that the telomere mutations found in LoVo cell line arose predominantly from defects in *MSH2* gene.

5.4.3 Telomere length in shMSH2 clones

Telomerase is a reverse transcriptase that adds *de novo* telomere repeats to the end of the telomeres and thus maintains telomere length. Most somatic cells show minimal or no telomerase activity and therefore enter cell cycle arrest (senescence) once their telomeres get to a critical length. One way to immortalise a cell line is by expression of the telomerase catalytic subunit, *hTERT*, resulting in telomerase activity. Cells transfected with *hTERT* were used for the *MSH2* inhibition experiments either by shRNA or by the use of $CdCl_2$.

Expression of a shRNA was used to obtain stable downregulation of *MSH2* expression in the CCD34-Lu/hTERT-1 cell line. In total 5 clones with different degrees of *MSH2* downregulation were generated and their Xp/Yp telomere length was determined using the STELA technique (Baird *et al.*, 2003). Telomere length of three shMSH2 expressing clones (clones 3, 4 and 5) was assayed at the Xp/Yp telomere and two of them (3 and 5) showed telomere shortening length of ~1 kb over 3 PDs (Figure 5.11) but clone 4 did not. The Xp/Yp telomere erosion rate was about 350 bp and 450 bp per PD in clones 3 and 5 respectively but the reported telomere erosion rate in normal human fibroblast is around 50 - 65 bp per PD (Counter *et al.*, 1992; Harley *et al.*, 1990). Both clones that showed telomere shortening also presented the lowest levels of *MSH2* expression and yet all three clones (3, 4 and 5) retained telomerase activity.

The unusual high erosion rate observed in clones 3 and 5 could be because of variable telomerase expression (though quantitative analysis suggests the level of expression is similar). However, it is also possible that the downregulation of *MSH2* contributed to telomere loss, in which case most telomere mutations caused by a lack of MMR would be deletions. Although there is not yet enough data from the shMSH2 experiments to be certain of the deletion bias, mutation analysis in LoVo cell line (*MSH2*$^{-/-}$) showed that there was an excess of deletions (five deletions out of seven mutations found, see table 3.4). To explore this preliminary finding, more clones with stable *MSH2* downregulation and controls need to be screened to determine telomere length and rate of shortening. In addition, more molecules need to be assayed for mutations in order to find out whether there is a bias towards deletions at telomeres.

Chapter 6

Extraordinary Instability in Human Telomeres Caused by Arrays of $(CTAGGG)_n$ Telomere Repeats

6.1 Introduction

Human telomeres are formed by several kilobases of tandem arrays of TTAGGG repeats but at the start, sequence variant-telomere repeats are interspersed with the consensus TTAGGG repeats. The determination of the distribution of variant repeats using TVR-PCR has shown that telomeres are highly variable however similarities between their interspersion patterns can be observed in samples that share the same haplotype in the telomere-adjacent region (Baird *et al.*, 1995; Baird *et al.*, 2000).

6.1.1 Telomere instability in the germ-line

Telomere mapping using TVR-PCR at the 12q and Xp/Yp telomeres has been carried out in families from a population of northern and western European descent (CEPH panel) in order to determine the telomere mutation frequency in the germ-line (Pickett, 2002). To obtain a mutation rate (at the 12q telomere) TVR-PCR was carried out in 160 children and their telomere maps were compared with those seen in the parents. One mutation was found in the CEPH family 1424 and consisted of a repeat-type change from a TTAGGG to a TCAGGG repeat together with an expansion of a block of TCAGGG repeats that involved approximately 8 repeat units. The mutation frequency at the 12q telomere was 0.62% (1/160; 95% CI, 0.03% – 3.6%).

Similar analysis at the Xp/Yp telomere identified two families (1362 and 1408) that intriguingly appeared to contain unusually high levels of germ-line instability at the haplotype B-associated Xp/Yp telomere. In family 1362, for example, three out of six children harboured mutations and three out of five children in family 1408 showed mutations (Pickett, 2002). Sequence analysis in these families showed that all mutations occurred at or immediately adjacent to a block of CTAGGG repeat, however no mutations were found in a further 165 children analysed lacking CTAGGG-containing telomeres. These results suggested that the CTAGGG blocks of repeats cause destabilisation at telomeres leading to a high mutation rate (Pickett, 2002).

6.1.2 Distribution of CTAGGG repeats within alleles

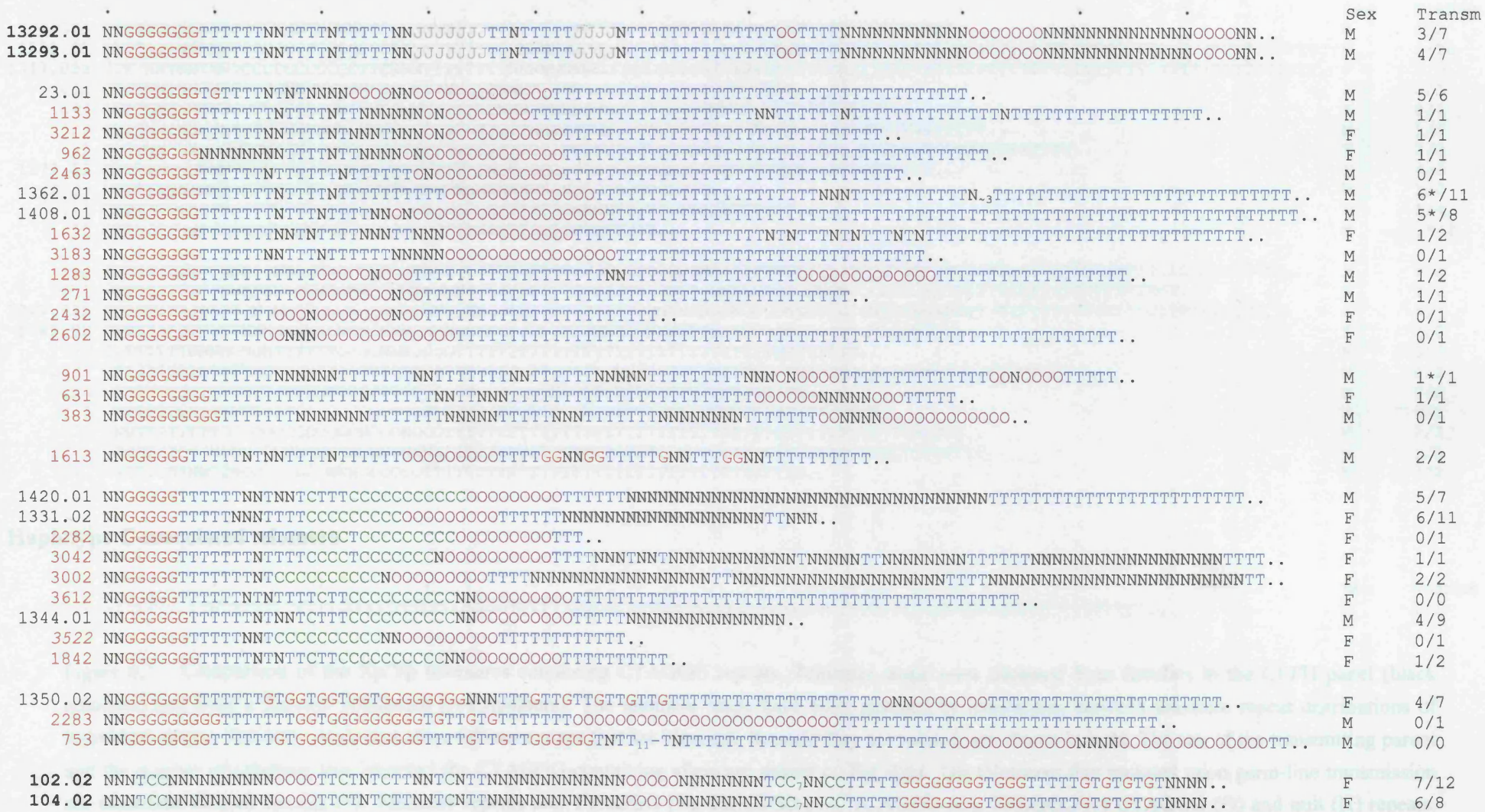
The prevalence of CTAGGG repeats was determined in a Russian population and in the parental DNAs that constitutes the CEPH panel at the Xp/Yp telomere by PCR using Xp/Yp specific primers (TS30A/T) and the telomere variant-repeat primer TAG-TelCTA2 to amplify blocks of $(CTAGGG)_n$ (Hills, 2004). Among the CEPH parents 10.6% of Xp/Yp alleles contained five or more CTAGGG repeats (17/160) whereas 5.8% of Xp/Yp alleles in the Russian population contained such telomere repeats (53/908). Thus overall ~6.6% (70/1068) Xp/Yp telomeres contained five or more CTAGGG repeats.

The samples with Xp/Yp CTAGGG-containing telomeres were amplified using the TVR-PCR technique to determine the number and distribution of $(CTAGGG)_n$. Comparison of the telomere maps (interspersion of TTAGGG and sequence variant repeats) showed they could be grouped according to their haplotype in the Xp/Yp adjacent sequence and arranged according to similarities between alleles (Figure 6.1). This revealed groups of related alleles that must have evolved along haploid lineages as described previously at the Xp/Yp and 12q telomere (Baird *et al.*, 1995; Baird *et al.*, 2000). The number and distribution of CTAGGG repeats varied between alleles, for example some samples contained single blocks of CTAGGG repeats composed of up to 25 repeats (*e.g.* sample 2283 from the Russian population), however other samples contained blocks of CTAGGG repeats that are interrupted by a few N-type repeats.

6.2 Aims

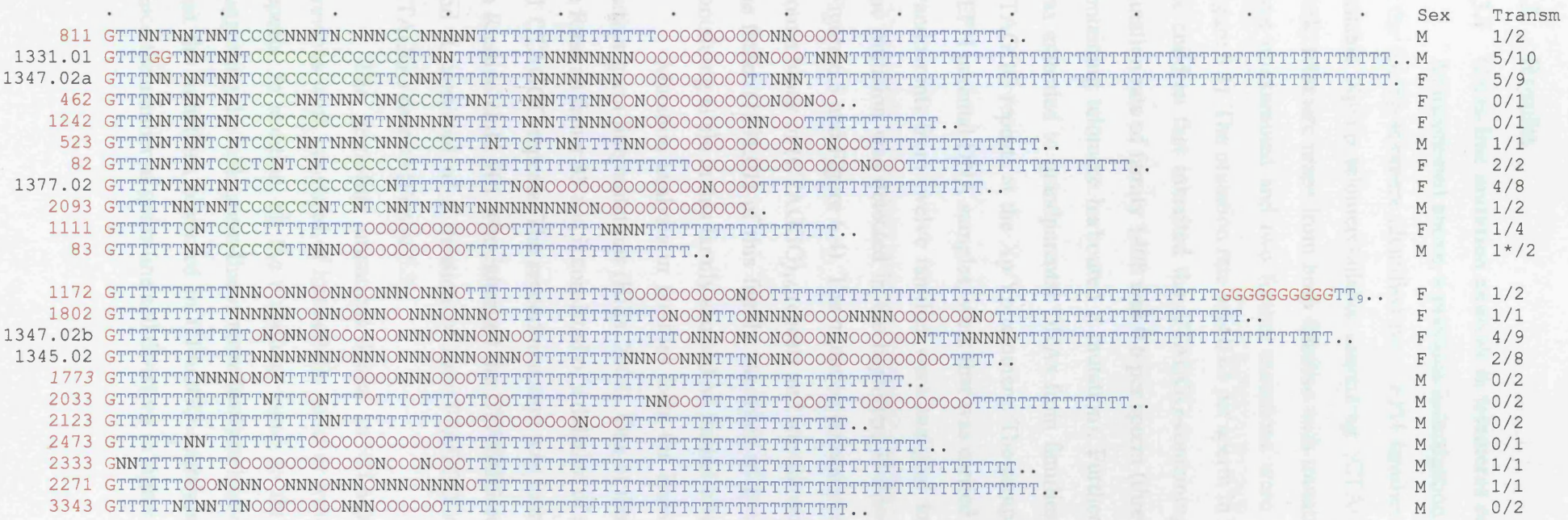
Germ-line mutation analysis revealed that some telomeres containing the variant CTAGGG-telomere repeats mutate at a high rate upon germ-line transmission compared to telomeres without such repeats, where only one mutation was found out of 325 alleles analysed (mutation frequency 0.3% per telomere per gamete) (Pickett, 2004 and Hills, 2006). In this work, the mutation analysis of telomeres containing CTAGGG repeats was extended to include the grandparents of a population of northern and western European descent (CEPH panel) samples at the Xp/Yp telomere in order to establish the mutation frequency and deduce the most likely mutation mechanism by comparing mutant maps. Furthermore, somatic samples from the colon were screened for the presence of such repeats and two colon DNAs were selected for mutation analysis to determine whether the instability observed in the germ-line is also present in a somatic tissue.

Haplotype A-associated telomere



Maps continue in next page…

Haplotype B-associated telomere



Haplotype C-associated telomere

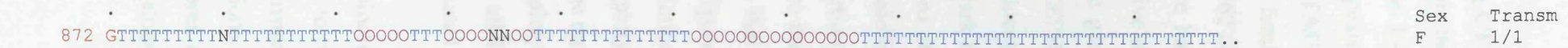


Figure 6.1 Comparison of the Xp/Yp telomeres containing CTAGGG repeats. Telomere maps were obtained from families in the CEPH panel (black numbers) and from a Russian population (red numbers). The telomere maps have been grouped by similarities between telomere repeat distributions of individual alleles. The individuals that share telomere maps that are identical, because they are related, are shown in bold. The sex of the transmitting parent and the number of offspring that inherited the CTAGGG-containing allele are shown on the right. The telomeres that mutated upon germ-line transmission are identified with an asterisk (*). Telomere repeats are, TTAGGG (T), TGAGGG (G), TCAGGG (C), TTGGGG (J) CTAGGG (O) and null (N) repeats. (This data was generated by M. Hills).

6.3 Results

6.3.1 Germ-line mutation analysis in telomeres containing CTAGGG repeats

As mentioned above, a previous investigation to estimate the germ-line mutation rate at the Xp/Yp telomere identified two CEPH families (1362 and 1408) that contained highly unstable Xp/Yp telomere-alleles containing $(CTAGGG)_n$ repeats (Pickett, 2002). In this work, telomere maps from both families with mutations in CTAGGG-containing telomeres were re-examined and two further mutations were found in family 1362 (Figure 6.2 and Figure 6.3). The mutation rate was 0.83 per sperm in family 1362 considering that five out of six children that inherited the CTAGGG-containing telomere harboured a mutation. The mutation rate of family 1408 was 0.6 per sperm (three out of five children with a CTAGGG-containing telomere harboured a mutation). Furthermore, the germ-line mutation analysis was extended to grandparental DNAs from families of the CEPH panel that contained the CTAGGG repeats at the Xp/Yp telomere. The samples were identified from the analysis of CEPH parental DNA samples, work that was carried out by Mark Hills (Figure 6.1). In total, grandparents from twelve families were analysed for mutations at the Xp/Yp telomere and one mutation was detected in the haplotype B-associated allele in the CEPH family 1331 (Figure 6.3 and Figure 6.4). The mutation consisted of the deletion of two CTAGGG repeats from a block of $(CTAGGG)_{16}$ repeats and was transmitted from the grandfather (1331.12) to the father (1331.01) of this family. Analysis of five children that inherited the CTAGGG-containing allele in this family showed no further instability (Figure 6.4).

Mutation analysis in families from the Russian population had also identified two further germ-line mutations (Hills, 2004). One of them in a haplotype B-associated telomere in Russian father 83 and it consisted of a deletion of one N-type repeat adjacent to a block of 11 CTAGGG repeats. The second mutation was found in a haplotype A-associated telomere in Russian child 901 who inherited the CTAGGG-containing allele from his/her father 903 and it consisted of a deletion of one TTAGGG repeat located proximal to a block of CTAGGG repeats (Figure 6.3).

The germ-line mutations found in CTAGGG-containing telomeres in this and previous work consisted of intra-allelic gains or losses of one to two CTAGGG or adjacent repeats. In addition, all the mutations found so far were transmitted from father to child suggesting the male germ-line is more likely to harbour mutations. Thus it was hypothesised that the mutations observed in CTAGGG-containing telomeres are replication dependent since the number of germ-line cell divisions is higher in spermatogenesis than in oogenesis.

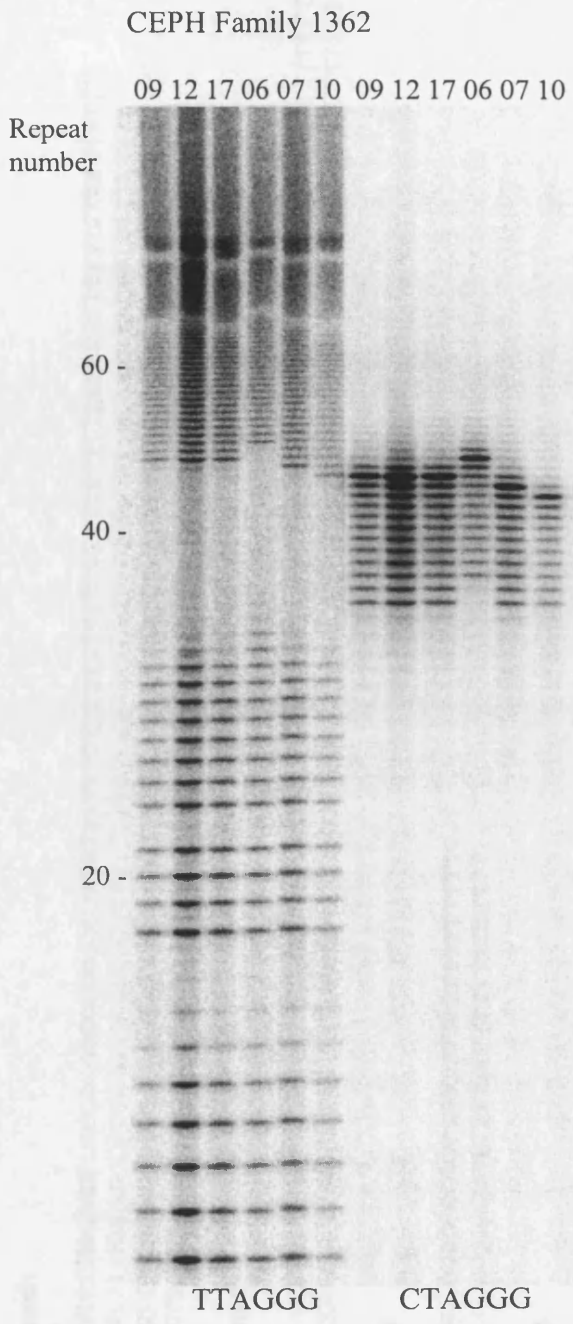


Figure 6.2 Germ-line telomere mutations found in the CEPH family 1362. Telomere maps from six children that inherited the Xp/Yp CTAGGG containing telomere are presented. Mutations in the CTAGGG or adjacent repeats were found in 1362.12 (daughter), 1362.17 (son), 1362.06 (daughter), 1362.07 (daughter) and 1362.10 (daughter) but not in the 1362.09 (daughter). The father (1362.01) from whom this allele was inherited is not shown.

Germ-line mutations at the Xp/Yp telomere

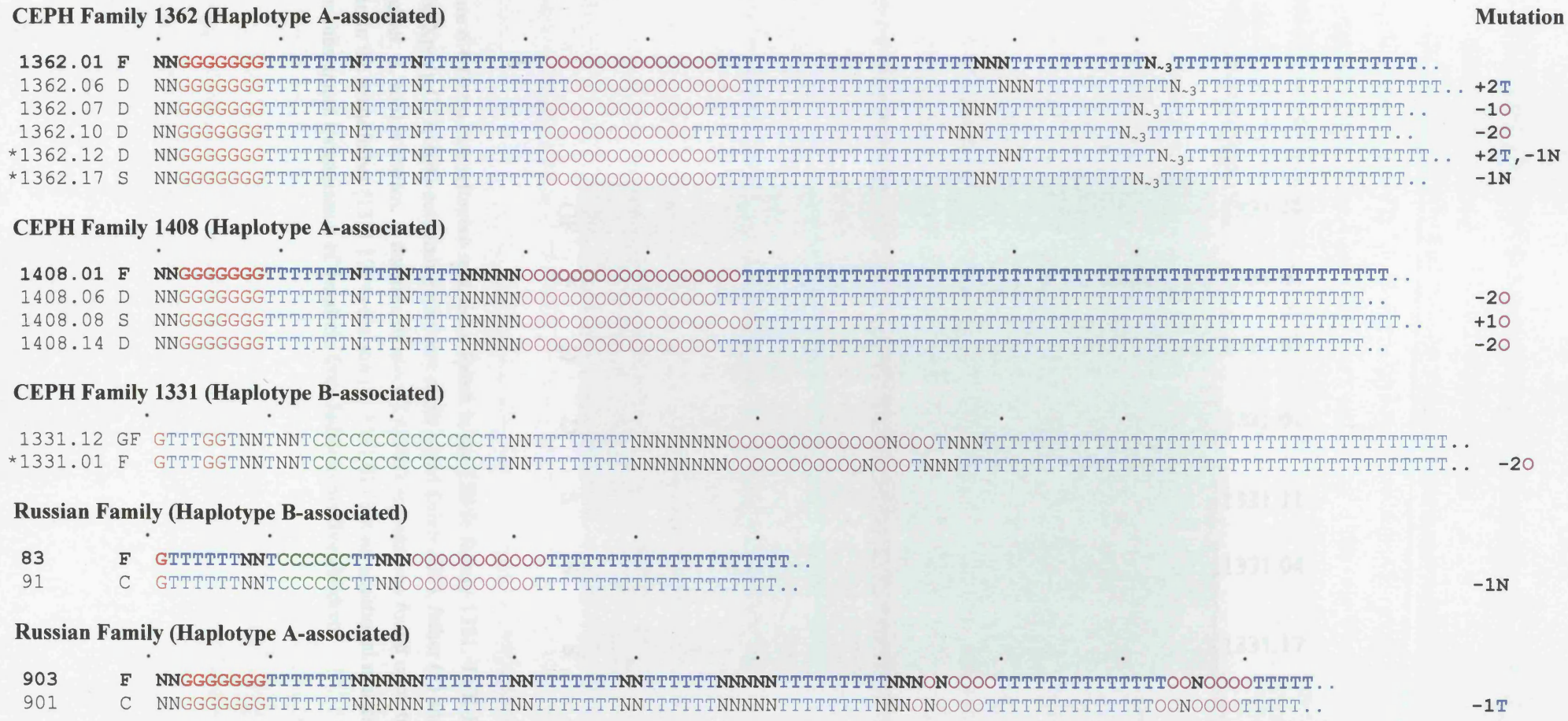


Figure 6.3 Germ-line mutations in families with Xp/Yp CTAGGG-containing telomeres. The interspersion of TTAGGG and sequence-variant repeats is shown for the parent (bold type) and children who showed mutations. The interspersion patterns of TTAGGG (T), TGAGGG (G), TCAGGG (C), TTGGGG (J) CTAGGG (O) and null (N) repeats are shown from the first repeat in the array (on the left). The mutations are summarised on the right *e.g* +2T indicates that the mutated allele in the child contained an additional two TTAGGG repeats. (*) mutations found in this work. Abbreviations are GF, grandfather, F, father; S, son; D, daughter, C, child.

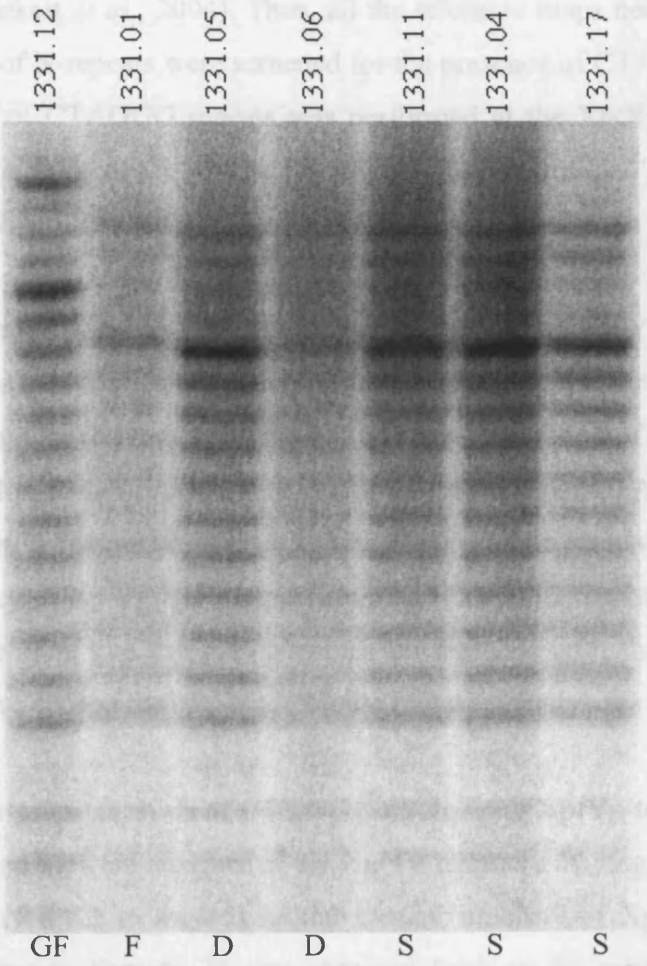


Figure 6.4 Germ-line telomere mutations found in the CEPH family 1331. TVR-PCR amplification of the Xp/Yp CTAGGG-containing telomere in the grand father (GF), father (F) and five children (D, daughter; S, son) is shown. A deletion of two CTAGGG repeats was found upon transmission of the allele in the grandfather (1331.12) to his son (F, 1331.01) but no additional mutations were detected upon subsequent transmission of the allele from father to his five children.

6.3.2 Instability of CTAGGG-containing telomeres in somatic cells

6.3.2.1 Selection of samples with CTAGGG repeats

DNAs from a panel of normal colon tissues were screened for the presence of CTAGGG repeats. The Xp/Yp flanking haplotype and telomere maps for all the samples in this panel had been determined previously but without the use of a primer that amplifies CTAGGG repeats (Pickett *et al.*, 2004). Thus, all the telomere maps determined previously that contained blocks of N-repeats were screened for the presence of CTAGGG repeats.

The detection of CTAGGG repeats was performed at the Xp/Yp telomere by PCR using the allele specific primers TS-30A or TS-30T and TelC (amplify TTAGGG repeats) or TAG-TelCTA2 (amplify CTAGGG repeats). The PCR reactions were cycled 23 times as follow: 96°C for 20 sec, 67°C for 40 sec and 70°C for 2 min. The PCR products were size-separated in 1% agarose gels, blotted and hybridised using an Xp/Yp telomere-specific probe. All the samples that appeared to have CTAGGG repeats were further amplified using the TVR-PCR technique to determine the number and the distribution of the repeats. In total, 34 alleles from 34 colon samples were screened for the presence of CTAGGG repeats and 4 samples were found to contain such repeats in one Xp/Yp allele (colon samples 11, 20, 23 and 26). The frequency of Xp/Yp telomeres containing CTAGGG repeats in the colon panel was 11.8% per allele. The two samples (samples 20 and 23) with the longest blocks of CTAGGG repeats were chosen for somatic telomere mutation analysis.

6.3.2.2 Somatic mutation analysis of CTAGGG containing Xp/Yp telomeres

Samples 20 and 23 were analysed at the Xp/Yp telomere by single molecule STELA (SM-STELA) and TVR-PCR to investigate the somatic stability of Xp/Yp telomeres that contain CTAGGG repeats. Sample 20 was obtained from an 81 year old female and it contained two blocks of CTAGGG repeats with the Xp/Yp haplotype B-associated telomere, the first one composed of 4 repeats and the second one formed by 7 repeats. In total, 102 molecules were analysed in this sample at the Xp/Yp haplotype B-associated telomere and 2 mutations were observed giving a telomere mutation frequency of 2% per telomere (95% CI, 0.1% – 7.3%; Figure 6.5). Mutant molecule 1 contained a repeat-type change of an N-type to a CTAGGG-type at repeat 13 from the start of the telomere. The second mutant molecule was more difficult to define because the resolution of the polyacrylamide gel after repeat 90 was poor. The mutation might have arisen by a large deletion that involved more than 60 telomere repeats distal to repeat 29 (Figure 6.5 and Figure 6.6).

Somatic mutations at the Xp/Yp telomere

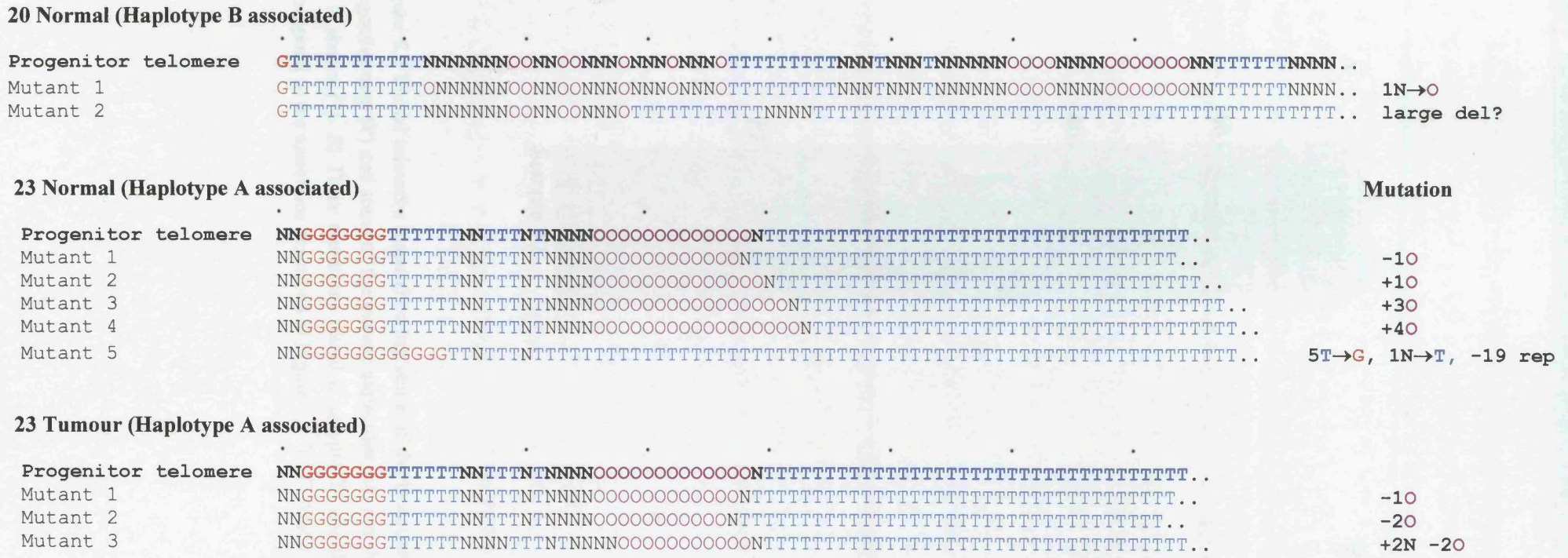


Figure 6.5 Somatic mutations found at Xp/Yp CTAGGG-containing telomeres in normal and tumour colon samples. The interspersion of TTAGGG and sequence-variant repeats along the array is shown for the progenitor telomere (bold type) and mutant molecules. The interspersion patterns of TTAGGG (T), TGAGGG (G), TCAGGG (C), TTGGGG (J) CTAGGG (O) and null (N) repeats are shown from the first repeat in the array (on the left). The mutations are summarised on the right *e.g* 5T→G indicates that five TTAGGG repeats had change to TGAGGG-type repeats in the mutant molecule. The mutation in the mutant molecule 2 (sample 20) was difficult to define but it might have arisen as a large deletion (intra-molecular) or by another inter-molecular mechanism.

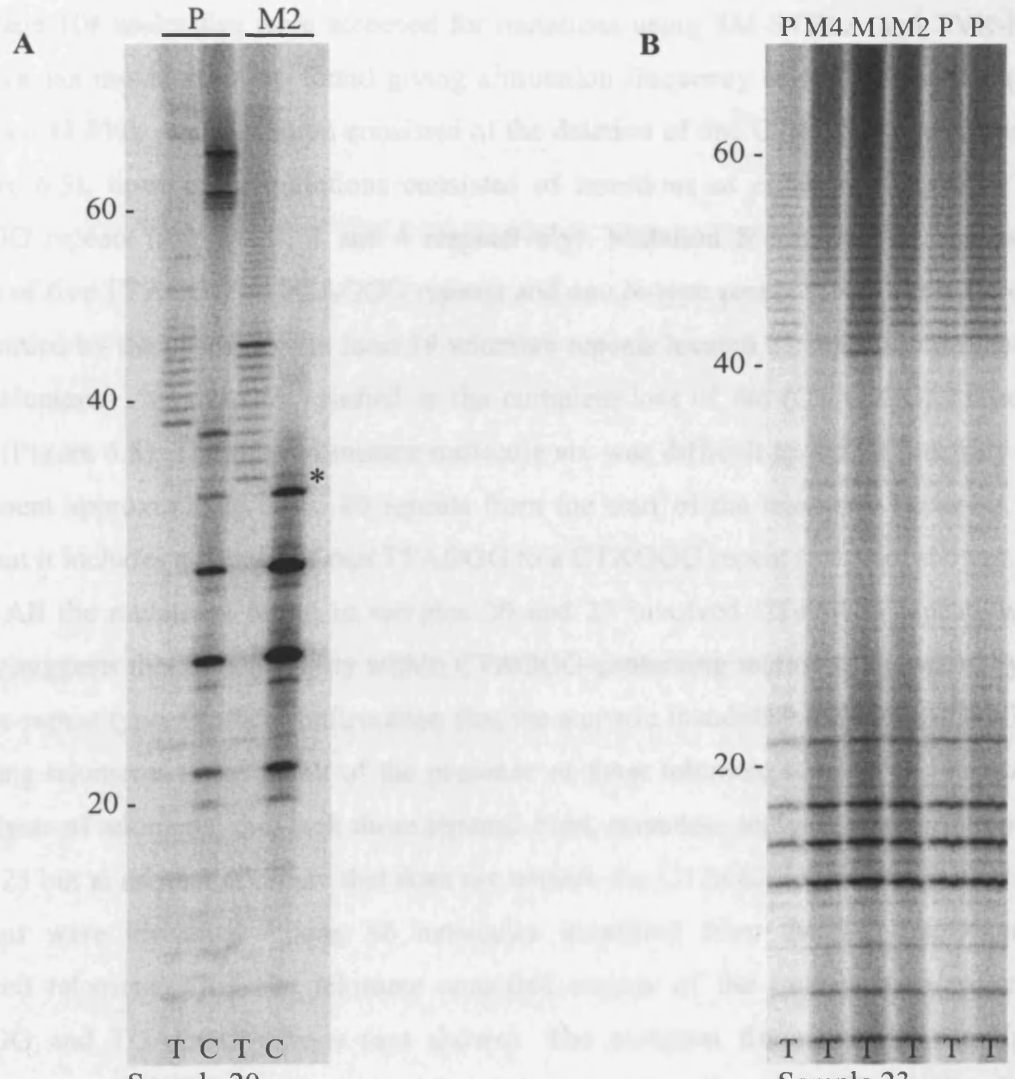


Figure 6.6 Partial telomere maps showing some of the mutations found in samples 20 and 23. The progenitor map (P) and some of the mutant molecules (M) are shown. *A*) Mutation 2 (M2) in sample 20 is shown (*). *B*) Three mutations found in sample 23 are shown. Mutant molecules 1, 2 and 4 correspond to the mutations described in Figure 6.5. T, TTAGGG; C, TCAGGG.

Sample 23 contains one block of 13 CTAGGG repeats located 26 telomere repeats from the start of the Xp/Yp haplotype A-associated telomere. Colon sample 23 was obtained from an 88 year-old female patient. Telomere mutation analysis was carried out in this sample and 108 molecules were screened for mutations using SM-STELA and TVR-PCR. Altogether six mutations were found giving a mutation frequency of 5.6% per allele (95% CI, 2.3% – 11.8%). One mutation consisted of the deletion of one CTAGGG repeat (mutant 1; Figure 6.5), three more mutations consisted of insertions of either one, three or four CTAGGG repeats (mutants 2, 3 and 4 respectively). Mutation 5 consisted of repeat-type changes of five TTAGGG to TGAGGG repeats and one N-type repeat to a TTAGGG repeat accompanied by the deletion of at least 19 telomere repeats located 21 repeats from the start of the telomere. This deletion resulted in the complete loss of the $(CTAGGG)_{13}$ block of repeats (Figure 6.5). The map of mutant molecule six was difficult to define precisely as it was present approximately 70 to 80 repeats from the start of the telomere. However, it is likely that it includes a change of one TTAGGG to a CTAGGG repeat (map not shown).

All the mutations found in samples 20 and 23 involved CTAGGG repeats which strongly suggests that the instability within CTAGGG-containing telomeres is caused by this telomere-repeat type. Further confirmation that the somatic instability of Xp/Yp CTAGGG-containing telomeres is the result of the presence of these telomere repeats was sought via the analysis of telomeres that lack these repeats. First, mutation analysis was carried out in sample 23 but at another telomere that does not contain the CTAGGG variant repeats and no mutations were identified among 86 molecules amplified from the 12q haplotype A-associated telomere. This 12q telomere consisted mainly of the interspersion pattern of TTAGGG and TGAGGG repeats (not shown). The mutation frequencies at the Xp/Yp CTAGGG containing telomere and the 12q telomere are significantly different ($P = 0.038$; Fisher's exact two-tailed test; Table 6.1). As a further control, a haplotype B associated Xp/Yp telomere that does not contain CTAGGG repeats in another normal colon sample (sample 2) was screened for mutations but none were found (Sample 2 was also described in Chapter 4). The combined somatic mutation frequency for the two CTAGGG containing telomeres (20 and 23 Xp/Yp telomere) is significantly different from the telomeres without such repeats (sample 2 at the Xp/Yp telomere and 23 at the 12q telomere) (Fisher's exact test, 2-tailed $P = 0.009$).

Sample	Sex/Age (yrs)	Telomere analysed (haplotype)	Number of CTAGGG repeats	No of mutations/ molecules analysed	Telomere mutation frequency (%)
23N	F/88	Xp/Yp (A)	13	6/108	5.6
23T	F/88	Xp/Yp (A)	13	3/98	3.1
23N	F/88	12q (A)	0	0/86	0
20N	F/81	Xp/Yp (B)	7[a]	2/102	2
2N	M/66	Xp/Yp (B)	0	0/92	0

Table 6.1 Somatic CTAGGG-containing telomere mutations in colon samples. [a]Sample 20N contains multiple CTAGGG repeats; the largest consecutive block contained 7 repeats. Haplotype of the telomere-adjacent sequence is given. Samples were obtained from normal colon (N) or from colon carcinoma (T).

Both patients 20 and 23 developed colorectal cancers, furthermore the tumour in patient 23 showed instability at a mononucletide (A_{10}) of the *TGFβRII* gene which is associated with MSI in 90% of colorectal cancers (Parsons *et al.*, 1995). To determine whether tumour formation in the presence of MSI further destabilised the Xp/Yp-CTAGGG containing telomere in this patient, the same telomere was investigated in the tumour sample (23T). In total, 98 molecules were screened for mutations and three mutations were found. Two of them consisted of deletions of one and two CTAGGG repeats (mutant alleles 1 and 2 according to Figure 6.5). Mutant allele 3 harboured two changes; the insertion of two N-repeats at repeat 16 (from the start of the telomere) and a deletion of two repeats from the block of $(CTAGGG)_{13}$ repeats. As seen in normal colon sample from patient 23, all the mutations were associated with the CTAGGG repeats. Furthermore there were no significant difference between the mutation frequencies in the tumour (23T) and normal colon samples (23N) from this patient ($P = 0.5$).

6.4 Discussion

6.4.1 Factors affecting the instability in CTAGGG-containing telomeres

Telomeres often contain an interspersion of the consensus TTAGGG telomere repeats with variant repeats (*e.g.* TGAGGG, TCAGGG, TTGGGG and CTAGGG) at the start of the array. Among these variant repeats, it was found that the presence of blocks of

more than five CTAGGG repeats often induces telomere instability in the germ-line and in somatic tissue. The mutation analysis carried out in CTAGGG-containing telomeres in the germ-line showed that all the mutations have occurred upon transmission of the allele from the father indicating that mutations in blocks of CTAGGG are more likely to occur in the male than the female germ-line. Nonetheless some blocks of CTAGGG repeats were transmitted from the father to multiple children without mutations, for example haplotype A-associated Xp/Yp telomere in CEPH father 23.01. This telomere contains $(CTAGGG)_{17}$ repeats but it is interrupted by two N-type repeats whereas the telomeres that showed mutations in CEPH fathers 1362.01 and 1408.01, that contained $(CTAGGG)_{13}$ and $(CTAGGG)_{18}$ blocks respectively, were without interruptions. Similarly, the haplotype B-associated telomere in the grandfather 1331.12 contained a block of $(CTAGGG)_{16}$ repeats interrupted by one N-type repeat and it mutated upon transmission to the father (1331.01). However the mutated allele containing $(CTAGGG)_{14}$ was stable upon further transmission to 5 children. Thus, it is likely that the size of the CTAGGG array and its homogeneity influence the stability of telomeres with such blocks of repeats. This has been observed in microsatellites, where a single CAT repeat is enough to stabilise an homogenous array of $(CAG)_{40}$ (Chong *et al.*, 1995). In addition, all the mutations whether in the germ-line or in somatic tissue had blocks of $(CTAGGG)_n$ repeats surrounded by few N-type repeats and/or TTAGGG repeats (Figure 6.3 and Figure 6.5) but no other telomere-variant repeats. Therefore it is likely that the surrounding repeats also play an important role in the stability of CTAGGG-containing telomeres.

6.4.2 Possible mechanisms leading to mutations in CTAGGG-containing telomeres

The mutant alleles in the germ-line harboured deletions or insertions of one to two repeats in blocks of $(CTAGGG)_n$ or adjacent repeats. This instability was also observed in two somatic samples derived from colon tissue. However, the mutations involved larger number of repeats. The mutations consisted of insertions of up to four telomere repeats and large deletions that could involve up to 60 telomere repeats in sample 20 and approximately 19 repeats in sample 23. This deletions could be the result of intra-allelic process such as t-loop excision or unequal telomere sister chromosome exchanges, however there is no direct evidence for this. In addition, a more complex process might have occurred such as a telomere-telomere recombination event (inter-allelic mechanism). Apart from these large deletions, all other mutations found in the germ-line and in somatic cells involved up to four

telomere repeats and can be explained by intra-allelic mechanisms such as replication slippage, gene conversion or unequal sister chromatid exchanges.

It is not clear what the cause of the high instability in CTAGGG-containing telomeres is, however it is possible that CTAGGG repeats form higher order DNA structures during replication promoting stalled replication forks and thus mutation events. For example, it has been shown that G-rich sequences such as minisatellites and telomeres can form antiparallel guanine quadruplex also known as G-quadruplexes *in vivo* (Paeschke *et al.*, 2005; Nakagama *et al.*, 2006). The formation of G-quadruplex at telomeres in ciliates depends on the binding of telomere proteins such as Tebpα and Tebpβ. However, during replication Tebpβ is phosphorylated leading to the dissociation of the Tebpα-Tebpβ complex from telomeres promoting the unfolding of G-quadruplex DNA structures (Paeschke *et al.*, 2005). The unfolding of DNA secondary structures is important to maintain the fidelity and stability of DNA during replication. Studies in *E. coli* have shown that structures known as G-loops are formed during replication of plasmids containing G-rich sequences (Duquette *et al.*, 2004). The formation of such structures in the template strand can potentially lead to deletions of repeats whereas insertion of telomere repeats may be present if secondary structures are formed in the newly synthesised strand. Furthermore, in human cells it has been found that different chemical compounds that stabilise G-quadruplexes can interfere with replication at telomeres (Riou *et al.*, 2002).

The formation of CTAGGG secondary structures at telomeres may also occur in similar variant repeats such as TCAGGG. However, instability at telomeres has only been associated to CTAGGG repeats. Thus is it possible that other factors may play important roles in the instability observed in CTAGGG such as the proteins involved in resolving secondary structures. It has been shown that heterogeneous nuclear ribonucleoproteins such as hnRNP A1 and hnRNP D bind to telomeres and can resolve G-quadruplex and G-G hairpin structures formed by telomere repeats (Enokizono *et al.*, 2005; Zhang *et al.*, 2006). However the affinity of such proteins for G-rich repeats varies according to their sequence, for example, the binding of hnRNP A1 proteins is 3 times weaker in $(TTAGGG)_5$ repeats compared to $(TCAGGG)_5$ repeats (Nakagama *et al.*, 2006). It can be hypothesised that CTAGGG repeats have reduced binding affinities for proteins that unfold G-quadruplex or other higher order DNA structures thus increasing telomere instability during replication (Figure 6.7).

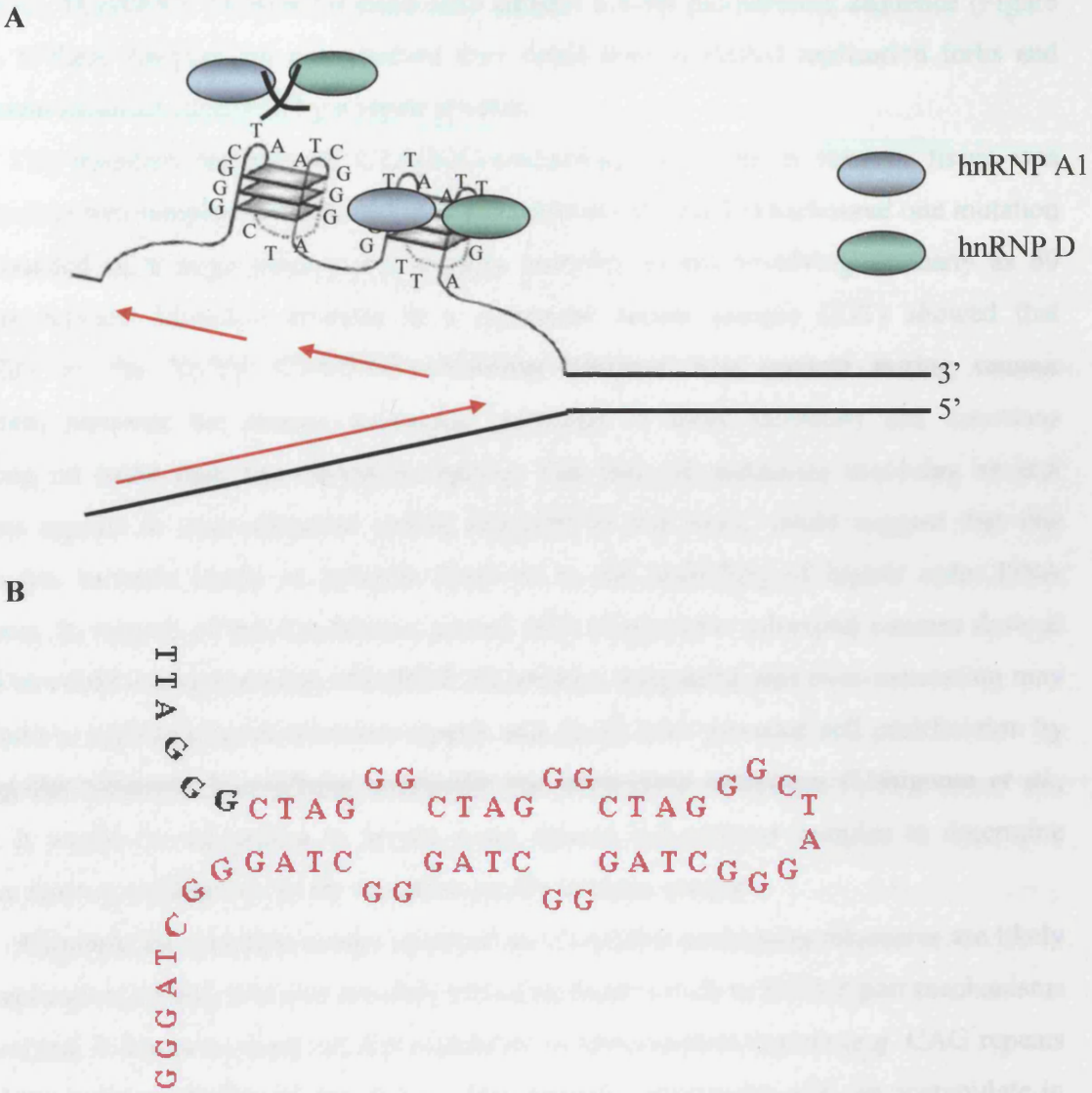


Figure 6.7 Model for CTAGGG repeat instability at telomeres. During replication telomere repeats can form DNA higher order structures such as G-quadruplex (A) or hairpin structures (B). Proteins such as hnRNA A1 and hnRNP D that unwind DNA secondary structure may have reduced affinity for CTAGGG repeats leading to mutations. Folded CTAGGG structures located in the template strand in replication would produce deletions of repeats (shown in figure) whereas folded structures in the nascent strand would produce insertion of repeats (not shown).

hnRNP A1 and hnRNP D are mentioned as examples of proteins that can affect the stability of CTAGGG repeats, however other proteins that resolve DNA secondary structures such as Bloom, Wrn, and Rtel (mouse homologue to *C. elegans* Dog-1) might also be involved (Ding *et al.*, 2004). In addition, CTAGGG repeats could potentially form hairpin structures that are thermodynamically more stable than other telomere repeats (*e.g.*

TTAGGG, TCAGGG, TGAGGG) since they contain a 4 bp palindromic sequence (Figure 6.7 B). If these hairpins are not resolved they could lead to stalled replication forks and subsequent mutation mediated by a repair process.

The mutation analysis of CTAGGG-containing telomeres in somatic tissue was performed in two samples from the colon. Both samples (20 and 23) harboured one mutation that consisted of a large deletion (or a more complex event) involving as many as 60 telomere repeats. Mutation analysis in a colorectal cancer sample (23T) showed that instability in the Xp/Yp CTAGGG-containing telomere was present during tumour formation; however the mutant molecules consisted of small deletions and insertions involving no more than two telomere repeats. The lack of mutations involving several telomere repeats in one colorectal cancer analysed in this work, could suggest that this sample has increase levels of proteins involved in the unfolding of higher order DNA structures. In support of this hypothesis, around 60% of sporadic colorectal cancers showed around two-fold over-expression of hnRNP A1 protein, suggesting this over-expression may contribute to maintenance of telomere repeats and could even promote cell proliferation by keeping the telomere 3'-overhang accessible for telomerase extension (Ushigome *et al.*, 2005). It would be interesting to screen more tissues and tumour samples to determine whether there is a difference in the mutation profile in these samples.

Although the mutation events observed in CTAGGG-containing telomeres are likely to be replication related, it is also possible that other factors such as DNA repair mechanisms are involved. It has been observed that instability in trinucleotide repeats (*e.g.* CAG repeats unstable in spinocerebellar ataxias) have a bias towards expansions and can accumulate in non-mitotic tissue such as the brain (reviewed in Pearson *et al.*, 2005). In addition, it has been observed that the expression of Msh2 protein in mice models of Huntington disease increased the instability of CAG repeats in somatic tissue (Manley *et al.*, 1999). These data suggest that MMR proteins can cause instability at trinucleotide repeats and in some cases this instability is not dependent on replication. Further investigation into the structures that CTAGGG repeats may form and the binding affinities for different proteins involved in the unfolding of higher order structures would be necessary to understand the molecular mechanism that lead to instability in CTAGGG-containing telomeres.

Chapter 7

Telomere Instability in Liposarcomas with Unknown Telomere Maintenance Mechanism

7.1 Introduction

7.1.1 Characteristics and classification of sarcomas

Sarcomas are mesoderm-derived tumours that affect connective tissue such as bone (osteosarcomas) or fat, muscle, tendons, synovial tissue and nerves (soft tissue sarcomas; STS). STS are rare tumours comprising around 1% of all malignancies in adults and are widely distributed in terms of anatomical location, however around 60% arise in the extremities. Many subtypes of STS have been described but the most commons are: malignant fibrous histiocytoma (28%), liposarcomas (15%), leiomyosarcomas (12%) and synovial sarcoma (10%) (Coindre *et al.*, 2001). Genetically, sarcomas can be divided into two main categories. In the first, sarcomas are characterised by having simple karyotypes with specific chromosome translocations while the second group is distinguished by having complex karyotypes and containing *TP53* mutations (Borden *et al.*, 2003). Analysis at telomeres in both groups have shown that sarcomas with complex chromosomal rearrangements often contain lengthened and heterogeneous telomeres compared to surrounding tissue as well as an increase in anaphase bridges. In addition, these tumours are often associated with activation of the ALT pathway as seen by the presence of ALT-associated PML bodies (APBs). On the other hand, telomeres from sarcomas with stable karyotypes contain similar or reduced telomere lengths and the tumours are often associated with activation of telomerase (Montgomery *et al.*, 2004).

7.1.2 Telomere maintenance mechanism in sarcomas

There are two main telomere maintenance mechanisms (TMM) in human cells: the activation of telomerase (TA) which is used by approximately 85% of human tumours (Kim *et al.*, 1994) and a recombination-base mechanism termed alternative lengthening of telomeres (ALT). In most sarcomas, contrary to other cancers, the proportion of tumours activating ALT is high and is similar to the proportion using TA. For example, in osteosarcomas it has been reported that only 44% use TA while 37% use the ALT pathway

to maintain their telomeres. Interestingly, 19% of the tumours analysed do not show signs of either the TA or ALT mechanisms and they have better survival prognosis (Ulaner *et al.*, 2003). In addition, osteosarcomas that use the ALT pathway to maintain their telomeres have increase end-to-end fusions and translocations/deletions as well as complex chromosome rearrangements indicating that in those tumours there is increased genomic instability (Scheel *et al.*, 2001).

Since the ALT pathway has been found to be frequent in sarcomas, some studies have determined the prevalence of this pathway in different types of sarcomas. For example, telomere heterogeneity and the presence of APBs, both hallmarks for ALT activation, were employed to determine the TMM status of 58 osteosarcomas, 50 astrocytomas and 101 STS (Henson *et al.*, 2005). The prevalence of ALT in these tumours was evident ranging from 47% in osteosarcomas to ~33% in both astrocytomas and STS. However within STS, the variation is large ranging from 77% using ALT in malignant fibrous histiocytomas to only 9% in synovial sarcomas.

7.1.3 Liposarcomas

As with other sarcomas, liposarcomas affect all age groups and represent 15% to 20% of all mesenchymal tumours (Sandberg, 2004). According to their cytogenetics and molecular characteristics, liposarcomas are divided in five main subtypes: myxoid, round-cell, de-differentiated, well-differentiated and pleomorphic liposarcomas.

Myxoid and round-cell liposarcomas are mostly found in the deep fat tissue of the extremities. These tumours are formed by stellar shaped cells with low grade malignancy and round shaped cells with high-grade metastatic potential respectively. Myxoid and round-cell liposarcomas are associated with few karyotype changes and in more than 90% of the cases a specific translocation between chromosomes 12 and 16, t(12;16)(q13;p11) is found (Knight *et al.*, 1995). Occasionally these tumours showed variant chromosomal alterations for example a translocation in chromosome 16, der(16)t(1;16), however they normally contain diploid karyotypes (Day *et al.*, 1997). Myxoid and round-cell liposarcomas often show high expression of *hTERT* which is associated with telomerase activity (Schneider-Stock *et al.*, 2003).

Well-differentiated (WD) and de-differentiated (DD) liposarcomas are the largest subgroup of liposarcomas accounting for approximately 45% of all liposarcomas. These tumours are formed by mature adipocytic cells with significant size variation. DD liposarcomas are similar to WD liposarcomas, although DD liposarcomas have an increased

metastatic rate compared to WD liposarcomas, thus it has been proposed that DD liposarcomas may progress from WD liposarcomas (Henricks *et al.*, 1997). However, both tumours have similar cytogenetic characteristics; for example, they normally contain small fragments of extrachromosomal DNA (double minutes), rings and giant chromosomes with a near-diploid karyotype (Pedeutour *et al.*, 1999).

Finally, pleomorphic liposarcomas are classified as high-grade aggressive sarcomas which are normally found in the extremities (Sandber, 2004). The cells of these tumours are large and contain multiple vacuoles; however the main characteristic of pleomorphic liposarcomas is that they contain multiple chromosome alterations as well as double minutes, rings and giant chromosomes (Mertens *et al.*, 1998). Apart from displaying complex chromosome rearrangements and polyploidy, it is common that pleomorphic liposarcomas have loss of heterozygosity and/or mutations of the *TP53* gene (Schneider-Stock *et al.*, 1998).

7.1.3.1 Prevalence of unknown TMM in liposarcomas

Liposarcomas are one of the best characterised STS in terms of the mechanism used for telomere maintenance. Liposarcomas show a high incidence of ALT activation compared to epithelial tumours, with a similar frequency of about 25% between the use of ALT and TA (Johnson *et al.*, 2005; Costa *et al.*, 2006). Interestingly, about 50% of the samples do not show any clear indication of ALT or TA. For example, tumours with heterogeneous telomere length lack often APB bodies while some tumours with short telomeres similar to TA, did not display telomerase activity. Furthermore, patients with these kinds of tumours were found to have a better survival prognosis (Costa *et al.*, 2006). Therefore, it has been proposed that liposarcomas with unknown TMM either use a mechanism not yet described or telomere length is not maintained (Johnson *et al.*, 2005).

7.1.4 Alternative mechanism for telomere maintenance

Telomere maintenance mechanisms other than TA or ALT have been described in some cell lines (Cerone *et al.*, 2005; Fasching *et al.*, 2005; Marciniak *et al.*, 2005). Similar to what has been seen in liposarcomas with unknown TMM, a SV40 immortalised fibroblast cell line (AG11395; derived from a patient with Werner syndrome), did not show telomerase activity or the conventional markers for ALT cells such as APBs. However, this cell line contains heterogeneous telomeres and extrachromosomal DNA structures (circular or linear) as seen in other ALT cells (Fasching *et al.*, 2005; Fasching *et al.*, 2007). Apart from the

absence of APBs, this cell line is different to the classical ALT cells since SV40 sequences contribute to the telomere maintenance. SV40 sequences are organised at telomeres in a tandem fashion, separated by telomeric repeats (Marciniak *et al.*, 2005). The organisation of these sequences at telomeres resemble the structure of type I survivors in yeast, where telomeres are elongated by amplification of Y' subtelomeric elements at the chromosome ends (Lundblad and Blackburn, 1993). In addition, it has also been reported that an immortalised clone (C3-cl6) derived from an ALT cell line (WI38 VA13/2RA) did not display the phenotype of an ALT or telomerase positive cell line (Cerone *et al.*, 2005). This clone retained some ALT features such as elevated T-SCE; however it did not have APBs or extrachromosomal DNA circles. Furthermore, C3-cl6 had short telomeres and it lacked telomerase activity thus is not clear how the telomeres were maintained. These two examples demonstrate that it is likely that different mechanisms for telomere maintenance exist, although they have not been described *in vivo*.

7.2 Aims

In order to maintain their telomeres, cells activate the enzyme telomerase or use a recombination-base mechanism (ALT). Evidence in sarcomas suggests that a large proportion of these cancers either use a different mechanism for telomere maintenance or they lack such a mechanism. By mapping the telomere variant repeat distribution at the start of telomeres, it was possible to distinguish cells that use the ALT pathway from the ones that used TA by the identification of recombination-based mutations at telomeres (Varley *et al.*, 2002). Furthermore instability at another locus, the MS32 minisatellite, is common in ALT samples suggesting an increase in recombination activity in those cells (Jeyapalan *et al.*, 2005).

In this work, the telomere stability was analysed in a group of liposarcomas that lack TA and markers for ALT to see if a TMM could be identified. For the first time in tumour samples, telomere-variant repeat analysis was carried out using a single molecule approach in order to identify mutations that arise at a low frequency and that are not visible using large-pool PCR amplifications. The instability at the MS32 minisatellite was also determined as a marker for increased recombination activity in the same group of liposarcomas (work conducted by Dr J Jeyapalan).

7.3 Results

7.3.1 Liposarcoma samples

Liposarcoma samples and information regarding telomere length and TMM status were provided by Dr. Nadia Zaffaroni (Instituto Nazionale Tumori, Italy) and are described in Costa *et al*, (2006). The telomere length and TMM of the liposarcomas used in this work are shown in table 7.1. Telomerase activity was determined using the TRAP assay (Wright *et al.*, 1995) and the percentage of APBs was used as a marker for ALT. APBs were identified by the colocalisation of promyelocytic leukaemia protein and telomeric DNA. A sample was considered ALT positive if it contained APBs in $\geq$ 0.5% of the cell analysed. In addition, telomere length was assayed by telomere restriction analysis (TRF), which consists in digesting genomic DNA with enzymes that do not recognise telomere repeats. The digested DNA is resolved in pulse field agarose gels and it is visualised by Southern blot hybridisation using a telomere probe.

7.3.2 Analysis of MS32 minisatellite stability in liposarcomas

MS32 is a GC-rich minisatellite located on chromosome 1q that is composed of tandem arrays of 29-bp units. The MS32 has a high germ-line mutation rate (0.81% per sperm) and the mutations found are generated by inter-allelic recombination events such as crossovers and conversions (Jeffreys *et al.*, 1994). In addition, somatic cells showed a mutation rate of $\leq$ 0.004% per haploid genome with mutations comprising simple deletions or insertions (Jefferys and Neumann, 1997). Instability at the MS32 minisatellite has been identified in ALT cell lines and in some ALT STS but not in telomerase positive samples. The mechanisms involving the mutations in ALT differ from the germ-line but they may include sister-chromatid exchanges. Thus, instability at MS32 can be used as a marker for ALT activation and it is thought to be an indicator of increased recombination activity in those cells (Jeyapalan *et al.*, 2005).

MS32 instability was assessed in liposarcoma samples with a known or unknown TMM by Dr Jennie Jeyapalan. The level of instability was determined using small-pool PCR and Poisson analysis was used to estimate the number of amplifiable molecules in each PCR (Jeffreys *et al.*, 1994). The results showed a higher instability in ALT liposarcomas (mutation frequency of 0.01 - 0.09 per molecule) compared to TA (0 – 0.005 per molecule).

Sample	TMM	Telomere length distribution (kb) [a]	Tumour grade and histology [b]	MS32 mutation frequency/molecule
1	TA	12.5 – 2.5 (5.6)	2-DD	0
15	TA	>23.0 – 12 (16.6)	2-MRC	0.005
20	TA	>23.0 – 12 (17.5)	3-MRC	nd
60	TA	10.0 – 3.7 (8.0)	1-Myxoid	0
72	TA	48.5 – 5.8 (25)	1-MRC	0
2	ALT	>23.1 – 2.2 (8.5)	2-Pleomorphic	nd
4	ALT	>23.1 – 2.3 (14.3)	3-Pleomorphic	0.09
41	ALT	>48.5 – 3 (5.8)	3-Pleomorphic	0.051
56	ALT	>23.1 – 1.1 (23.0)	3-DD	0.025
63	ALT	30 – 3.7 (12.5)	3-Pleomorphic	0.063
66	ALT	>48.5 – 17 (10.4)	1-DD	0.01
12	**Unknown**	**16.5 – 2.5 (7.5)**	**1-DD**	**0.026**
14	Unknown	>23.1 – 5.0 (8.0)	2-DD	0
17	Unknown	12.5 – 3.8 (8.0)	1-Myxoid	0.006
30	Unknown	10.1 – 5.4 (7)	1-WD	0
32	**Unknown**	**>48.5 – 3.0 (11.4)**	**2-DD**	**0.039**
35	Unknown	10.1 – 5.2 (7.6)	1-WD	0.003
36	Unknown	9.5 – 5.8 (7.6)	1-WD	0
38	Unknown	7 -3 (4.4)	1-WD	0
42	Unknown	23.1 – 4.5 (8.4)	3-DD	0.008
43	Unknown	11 – 5.8 (9.5)	1-Myxoid	0
47	Unknown	10.1 – 5.4 (8.0)	1-WD	0.016
48	Unknown	11.0 – 4.2 (8.0)	3-DD	0
49	Unknown	10.4 – 5.8 (9.5)	3-DD	0.003
53	Unknown	18 -4.4 (11)	2-DD	0.003
58	**Unknown**	**11.0 – 5.8 (7.6)**	**1-WD**	**0.032**
67	Unknown	17 – 3.8 (9)	1-Myxoid	0.01
71	Unknown	48.5 – 2.5 (9.4)	1-DD	0

Table 7.1 Liposarcoma samples using the TA, ALT or unknown TMM. TMM was assayed in Costa *et al.* (2006) by determining activation of telomerase or APB status. [a]Telomere length distribution was obtained from Nadia Zaffaroni and the mean it is shown in brackets. Some liposarcomas with an unknown TMM showed MS32 instability to the levels seen in ALT tumours (shown in bold). [b]Histology grade abbreviations: DD, de-differentiated; MRC, myxoid round cells; WD, well differentiated.

The lower limit for an ALT phenotype was set as a mutation frequency of 0.02 per molecule. The level of instability in MS32 minisatellite varied between samples using the ALT pathway and the levels of instability correlated with the percentage of APBs in the cells. For example, sample 66 (Table 7.1) has a MS32 mutation frequency of 0.01 and 7% of the cells contained APBs whereas sample 63 had a MS32 mutation frequency of 0.063 and 43% of APBs.

MS32 instability was used as a marker for increased recombination activity in liposarcomas with unknown TMM and so the MS32 mutation frequency was determined (Figure 7.1; Table 7.1). The mutation frequencies varied from 0 – 0.039 per molecule in the 17 samples analysed. Three of the unknown TMM liposarcomas (sample 12, 32 and 58) showed MS32 instability above the 0.02 mutation frequency threshold. This suggests that these liposarcomas have an increase recombination-like activity and thus an ALT-like mechanism may be active despite the absence of APBs.

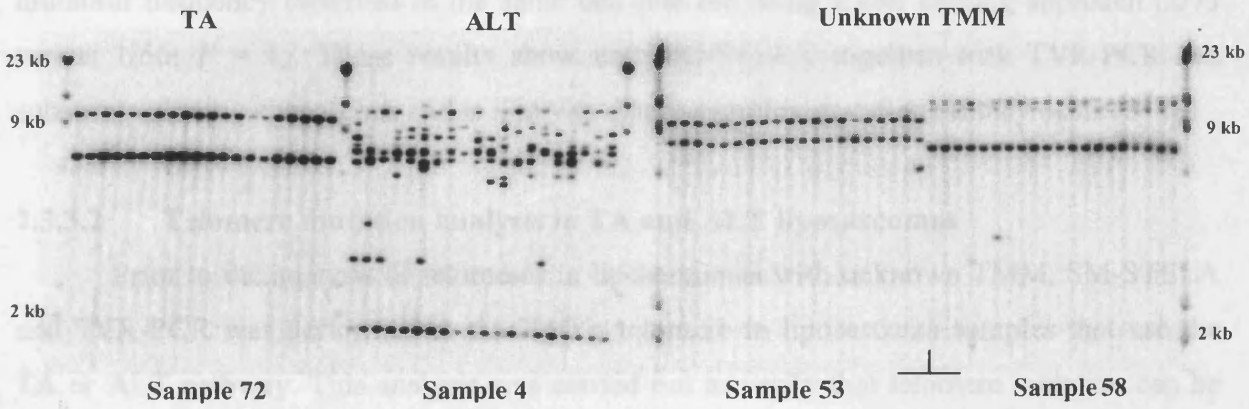


Figure 7.1 Small-pool PCR amplification of MS32 in liposarcomas using ALT, telomerase (TA) or an unknown TMM. Sample 72 (TA) and 53 (unknown TMM) show MS32 stability whereas sample 4 and sample 58 show MS32 instability with mutation frequencies of 0.09 and 0.032 per molecule respectively (Figure provided by Dr J Jeyapalan).

7.3.3 Telomere analysis

7.3.3.1 Detection of complex mutations using a single molecule approach

Analysis of telomeres using the TVR-PCR technique showed that ALT cell lines, but not telomerase positive or primary cell lines, undergo complex mutations that appear to be based on a recombination-like process. Such mutations occurred at low frequency and were

detected using cell cloning approaches (Varley *et al.*, 2002). Generally these mutations cannot be explained by simple insertion or deletions of telomere repeats, instead they involve the truncation of the progenitor telomere and addition of a novel telomere-repeat array. In order to assess whether this kind of mutation can be detected by using single molecule STELA (SM-STELA) together with TVR-PCR, an ALT cell line (IIIcf/a2 post-crisis) was used as a control. This cell line has been cloned previously and the mutation frequency of complex mutations was 2.7% (2/73) per molecule at the 16p/q telomere (Varley *et al.*, 2002).

In this work, SM-STELA at the 16p/q telomere was carried out on bulk DNA from the IIIcf/a2 post-crisis cell line using the allele specific primer Nitu14eC as described in 2.2.7.3 and 56 molecules were mapped by TVR-PCR. One complex mutation was found (1/56) at the start of the telomere comprising a change from a homogeneous TTAGGG array seen in the progenitor allele to an interspersion pattern of TTAGGG and null repeats (Figure 7.2). The mutation frequency was 1.8% per molecule and was statistically similar to the mutation frequency observed in the same cell line but using a cell cloning approach (2/73 versus 1/56; $P = 1$). These results show that SM-STELA together with TVR-PCR can substitute cloning approaches and is likely to detect complex mutations if they occur.

7.3.3.2 Telomere mutation analysis in TA and ALT liposarcomas

Prior to the analysis of telomeres in liposarcomas with unknown TMM, SM-STELA and TVR-PCR was performed at the Xp/Yp telomere in liposarcoma samples that use the TA or ALT pathway. This analysis was carried out to verify that telomere mutation can be detected *in vivo* using the single molecule approach described previously in cell lines.

Two samples (20 and 60) that use telomerase as their TMM were analysed using SM-STELA and TVR. In total, 182 and 49 Xp/Yp telomeres were analysed respectively and one mutation was found in sample 20. The mutation frequency was 0.5% per molecule (1/182) (95% CI, 0.01% - 3%). The mutant molecule harboured two simple mutations in the Xp/Yp haplotype A-associated telomere that consisted of repeat-type changes from null repeats (repeats that do not amplify with the primers used) to TTAGGG repeats at positions 16 and 21 from the start of the telomere. This is a simple mutation as it can be explained by intra-allelic mechanisms such as replication slippage, unequal sister chromatid exchange or a gene conversion-like process. The same analysis was carried out at the Xp/Yp haplotype B-associated telomere of sample 60; however no mutations were observed (0/49).

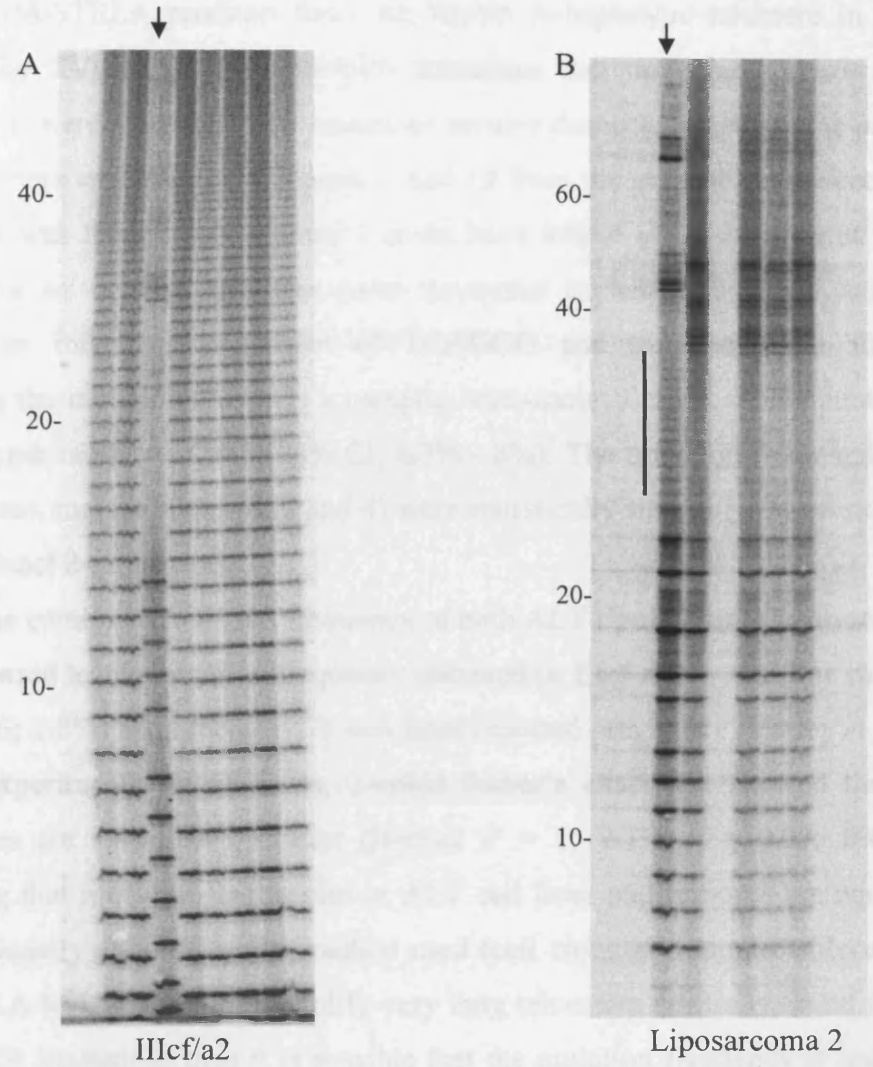


Figure 7.2 Complex mutations found in IIIcf/a2 cell line and in a liposarcoma sample that used the ALT pathway to maintain their telomeres. TVR-PCR was performed in a cell line (IIIcf/a2) (*A*) and in liposarcoma sample 2 (*B*) using SM-STELA products. Only amplifications of the TTAGGG repeats are shown. For each sample, the gel shows a collection of telomere maps from SM-STELA products and a mutant molecule is indicated with an arrow. Mutant molecule in IIIcf/a2 post-crisis has a fusion point at the start of the telomere repeat array while in liposarcoma 2, the fusion point occurs between repeats 25 - 38 (indicated with bar).

In addition, two liposarcomas samples (2 and 4) that use the ALT pathway, were screened for mutations at the Xp/Yp telomere. In total 94 SM-STELA products were screened by TVR-PCR at the Xp/Yp A-haplotype telomere from sample 2 and one mutation was found (1/94). The mutation frequency was 1.1% per molecule (95% CI, 0.03% - 5.7%). The mutant telomere harboured a complex mutation consisting of the replacement of the

progenitor telomere map with a novel telomere array at repeat 48 (Figure 7.2). In the same way, 88 SM-STELA products from the Xp/Yp A-haplotype telomere in sample 4 were analysed by TVR-PCR. Two complex mutations that must have arisen by inter-allelic mechanisms were identified. The mutations involve the replacement of the progenitor with a novel telomere repeat array at repeats 2 and 12 from the start of the telomere (Figure 7.3). Initially it was though that mutant 1 could have arisen *via* a deletion at the start of the telomere or by unequal telomere sister-chromatid exchange; however sequence analysis revealed an interspersion pattern of TCAGGG and the consensus TTAGGG repeat suggesting the mutation arose via a complex inter-molecular event. The mutation frequency was 2.3% per molecule (2/88; 95% CI, 0.3% - 8%). The mutation frequencies of both ALT liposarcomas analysed (sample 2 and 4) were statistically similar (1/94 *versus* 2/88 $P = 0.61$; Fisher's Exact 2-tailed test).

The combined mutation frequency of both ALT liposarcomas analysed (3/182; 1.6%) was compared to the mutation frequency obtained in IIIcf ALT+ cell line carried out in this work (1/56; 1.8%) and others ALT+ cell lines reported previously (Varley *et al.*, 2002) using cloning experiments. In all cases, 2-tailed Fisher's exact test showed that the mutation frequencies are statistically similar (IIIcf/a2 $P = 1$; WI38 $P = 0.07$; JFCF6 $P = 0.43$) suggesting that mutation frequencies in ALT cell lines and tumours are equivalent and are not significantly affected by the method used (cell cloning or single molecule approaches). The STELA technique cannot amplify very long telomeres normally found in ALT samples due to PCR limitations thus it is possible that the mutation frequency is underestimated by using the STELA technique and TVR-PCR if mutations are present in large telomeres however the results did not show a significant effect.

7.3.3.3 Telomere mutation analysis in unknown TMM liposarcomas

Telomeres from liposarcomas with unknown TMM were screened for mutations using SM-STELA and TVR-PCR at the Xp/Yp telomere. Around 100 molecules were analysed in five different samples (samples 12, 17, 32, 47 and 58), three of them displaying MS32 instability (Table 7.2). No mutant telomeres were identified in samples 12, 17 and 47 even though one of them, sample 12, showed MS32 instability. However complex mutations were found in samples 32 and 58 which were similar to the ones observed in ALT tumours. In addition, these two samples showed MS32 minisatellite instability.

Haplotype A-associated telomere

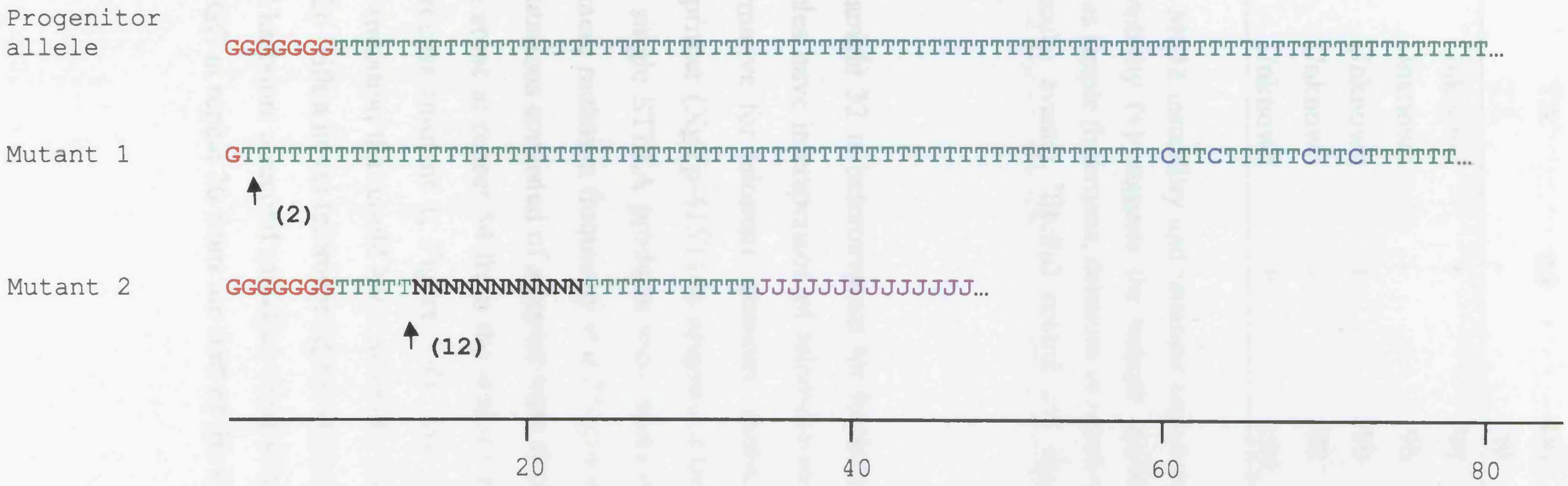


Figure 7.3 Telomere mutations found at the Xp/Yp telomere of liposarcoma 4

Telomere maps of the progenitor allele and two mutant molecules found in a liposarcoma (sample 4) that uses the ALT pathway. Mutant 1 harbours a complex mutation with a fusion point at telomere repeat 2 (indicated with an arrow). Mutant 2 also contains a complex mutation with a fusion point at telomere repeat 12. Both mutations consisted of the replacement of the progenitor telomere with a novel telomere. T, TTAGGG; G, TGAGGG; C, TCAGGG; J, TTGGG; N, repeats that do not amplify with the primers used.

Sample	TMM	MS32 instability	Molecules analysed	Type of mutations		Telomere mutation frequency per molecule (%)
				Simple [a]	Complex [b]	
IIIcf/a2 [c]	ALT	+	56	0	1	1.8
2	ALT	nd	94	0	1	1.1
4 *	ALT	+	88	0	2	2.3
20 *	TA	nd	182	1	0	0.5
60 *	TA	-	49	0	0	0
12	Unknown	+	98	0	0	0
17	Unknown	-	98	0	0	0
32	Unknown	+	99	1	1	2
47	Unknown	-	98	0	0	0
58	Unknown	+	104	1	1	1.9

Table 7.2 MS32 instability and telomere mutations at the Xp/Yp telomere among some liposarcomas. MS32 instability (+) suggests the tumour shows increased recombination activity. Mutations were classified as simple (insertions, deletions or repeat-type changes) or complex (mutations not explained by intra-molecular events). [c]IIIcf/a2 control cell line. nd, not determined. * Samples analysed by Dr J Jeyapalan.

Sample 32 is heterozygous for haplotypes A and B in the Xp/Yp subtelomeric region. Both alleles have interspersion of telomere-variant repeats at the start of the telomere and thus are informative for telomere mutation analysis. SM-STELA was carried out with the allele-specific primer (XpYp-415T) to amplify telomeres at the haplotype B-associated telomere. In total, 99 single STELA products were analysed by TVR-PCR and 2 mutations were detected. The telomere mutation frequency was 2% per molecule (95% CI, 0.25% - 7.1%; Table 7.2). One of the mutations consisted of a repeat-type change from a TTAGGG to a TGAGGG repeat. This mutation arose at repeat 54 from the start of the telomere and most probably arose by an intra-allelic process (mutant 1; Figure 7.4). The second mutation found in liposarcoma 32 was a complex mutation that could have arisen by the fusion of the progenitor telomere map at repeat number 26 with a novel telomere repeat array (mutant 2; Figure 7.4 and Figure 7.5). In addition, mutant 2 harbours a second mutation consisting of a repeat-type change of a TTAGGG repeat to a TGAGGG at repeat 20 from the start of the telomere (Figure 7.5).

Haplotype A-associated telomere

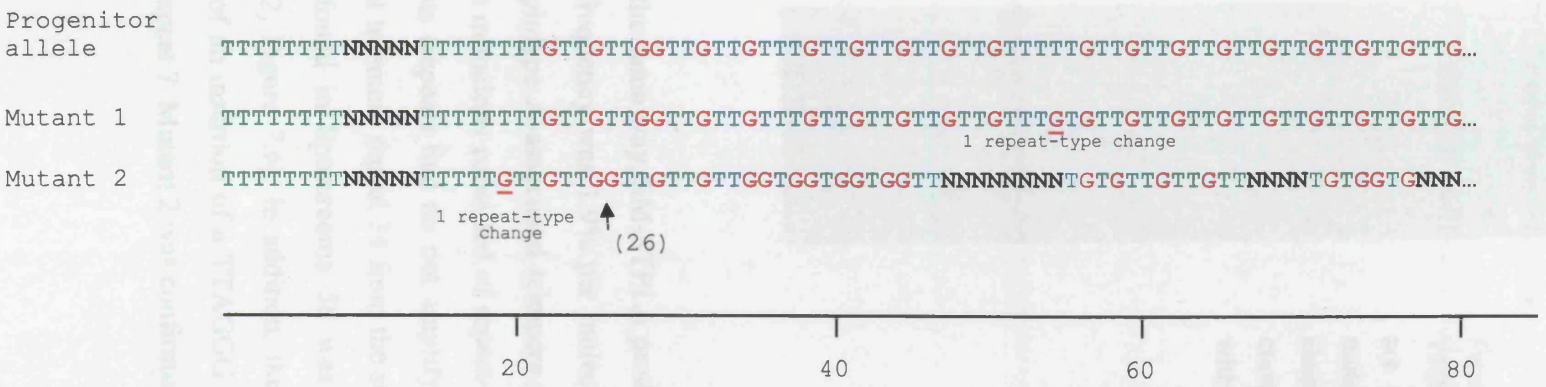


Figure 7.4 Telomere mutations at the Xp/Yp telomere of liposarcoma 32 with unknown TMM

Telomere maps of the progenitor allele and 2 mutant molecules of liposarcoma 32 are shown. Mutant 1 harbours a simple mutation involving the change of a TTAGGG repeat to a TGAGGG repeat. Mutant 2 comprises a change of a TTAGGG repeat to a TGAGGG repeat (underlined) and a fusion of the progenitor telomere map to a novel telomere map at repeat 26 (indicated with an arrow). T, TTAGGG; G, TGAGGG; C, TCAGGG; J, TTGGG G; N, repeats that do not amplify with the primers used.

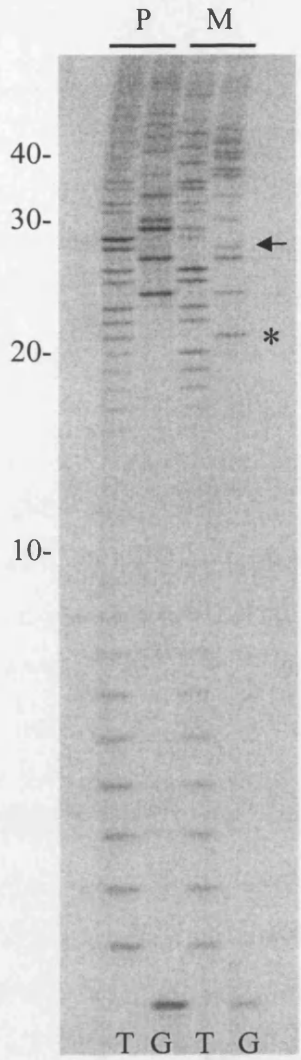


Figure 7.5 TVR-PCR of liposarcoma 32 with unknown TMM. The progenitor (P) and a mutant (M) telomere maps are shown. The mutant molecule harboured two different mutations, a simple mutation that consists of a repeat-type change from a TTAGGG to a TGAGGG repeat (*) and a complex mutation with fusion point at repeat 26 (indicated with an arrow). T, TTAGGG; G, TGAGGG.

In the same way, SM-STELA products from sample 58 were screened for mutations. The mutation frequency was 1.9% per molecule (95% CI, 0.23% - 6.8%; Table 7.2). In total 104 Xp/Yp haplotype A-associated telomere products were analysed and 2 mutations were found. One of the mutations consisted of repeat-type changes of four TTAGGG telomere repeat to N-type repeats (repeats that do not amplify with the primers used in TVR-PCR). This mutation occurred at telomere repeat 34 from the start of the telomere (mutant 1; Figure 7.6). The second mutation found in liposarcoma 58 was complex with a fusion point at repeat 28 (mutant molecule 2; Figure 7.6). In addition, this mutant molecule harboured a second mutation that consisted of an insertion of a TTAGGG repeat located close to the start of the telomere, after telomere repeat 7. Mutant 2 was confirmed by sequence analysis (Appendix IV).

Haplotype A-associated telomere

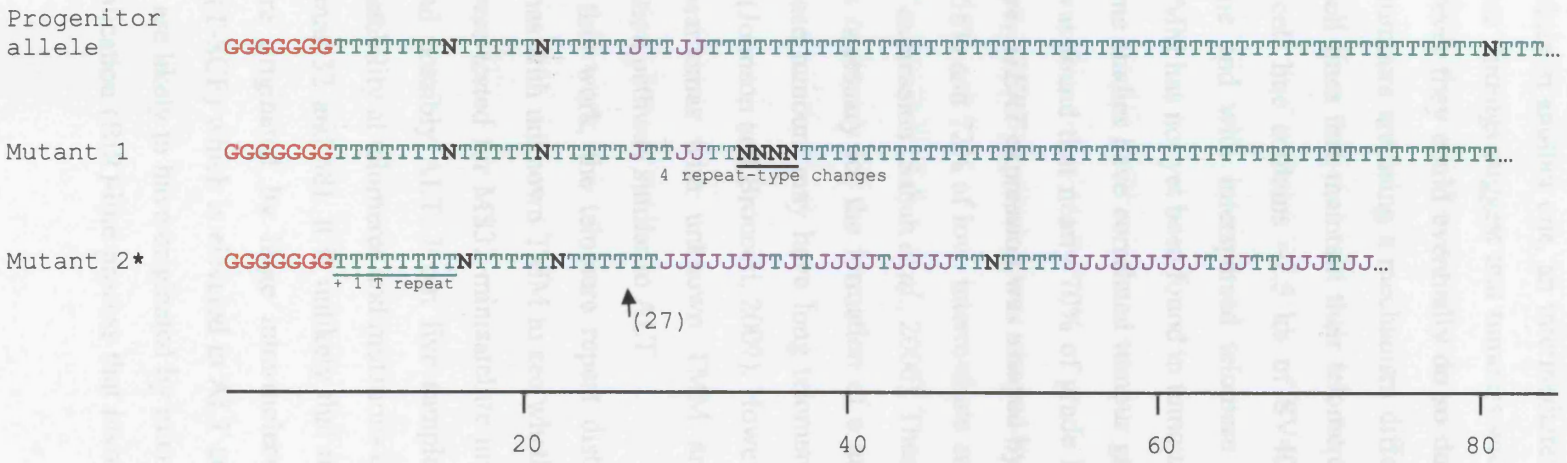


Figure 7.6 Telomere mutations at the Xp/Yp telomere of liposarcoma 58 with unknown TMM

Telomere maps of the progenitor allele and 2 mutant molecules in liposarcoma 58 are shown. Mutant 1 harbours a simple mutation involving the change of four TTAGGG repeats to N-type repeats (underlined). Mutant 2 comprises the insertion of one TTAGGG type repeat (underlined) and fusion of the progenitor to a novel telomere map at repeat 27 (indicated with an arrow). * Mutant 2 was confirmed by sequence analysis. T, TTAGGG; G, TGAGGG; C, TCAGGG; J, TTGGGG; N, repeats that do not amplify with the primers used.

7.4 Discussion

A large proportion of mesenchymal-origin tumours such as liposarcomas use the ALT pathway to maintain their telomeres. More interestingly, almost half of liposarcomas do not use a known TMM or they lack such a mechanism. However, the analysis of two recurrent liposarcoma tumours with an apparent lack of TMM, showed in one case that telomerase was activated while in another one, an intermediated APB phenotype was observed (Johnson *et al.*, 2005). These findings suggest that tumours with unknown TMM may not have yet activated a TMM however they could eventually do so during tumour progression. Another possibility is that these tumours are using a mechanism different to the TA or ALT. In support of this idea, there are cell lines that maintain their telomeres using different mechanism. For example, the AG11395 cell line contains ~2.5 kb of SV40 sequences organised in tandem arrays at the chromosome end with interspersed telomere repeats (Marciniak *et al.*, 2005). However, a different TMM has not yet been found in tumours.

Some studies have correlated tumour grade with TMM in liposarcomas. In Costa *et al.* (2006) it was found that nearly 70% of grade I tumours do not show evidence of ALT or TA. Furthermore, *hTERT* expression was assayed by immunohistochemistry in STS and it was found that 46%, 58% and 72% of low, intermediate and high-grade tumours respectively were positive for *hTERT* expression (Sabah *et al.*, 2006). These studies suggest that the activation of a TMM is not always necessary for the formation of a tumour. Moreover, it has been suggested that the cells of these tumours may have long telomeres thus, avoiding telomere crisis during tumour formation (Johnson and Broccoli, 2007). However, the results obtained in this work suggest that some liposarcomas with unknown TMM are indeed using a TMM that is based on a recombination pathway similar to ALT.

In this work, the telomere repeat distribution (of the first 500 bp) was analysed in liposarcomas with unknown TMM to see whether a TMM could be identified. Furthermore, the samples were tested for MS32 minisatellite instability as a marker of increased recombination activity and possibly ALT. From five samples with unknown TMM analysed, two of them showed instability at telomeres and mutations consistent with the activation of an ALT pathway (liposarcomas 32 and 58). It is unlikely that the complex mutations found in liposarcomas 32 and 58 are originated by large intra-molecular deletions or by telomere sister-chromatid exchange (T-SCE) which is elevated in ALT cells (Londono-Vallejo *et al.*, 2004). Instead these mutations are likely to have originated by inter-molecular recombination process such as break-induce replication (BIR)-like process that involves the repair of double strand breaks or stalled

replication forks by a recombination-dependent replication process (McEachern and Haber, 2006).

In addition, liposarcomas 32 and 58 have MS32 instability to the levels seen in ALT positive samples. This finding also suggests that these tumours are using a TMM similar to ALT. Interestingly, these tumours do not have APBs, thus suggesting that APBs are not necessary for telomere maintenance in these liposarcamas even though these structures has been reported to be essential in ALT cell lines (Jiang *et al.*, 2005). For example, over-expression of Sp100 protein, a constituent of APBs, has the ability to sequester APB recombination proteins such as Mre11, Rad50 and Nbs1. As a consequence, the ALT pathway is inhibited as seen by a decrease in telomere length and suppression of APB formation.

The two complex mutations found in the unknown TMM liposarcomas 32 and 58 resembled the complex mutations seen in ALT liposarcomas (sample 2 and 4), though they have some differences. For example, the novel repeat array added to the truncated progenitor molecules contained variant repeats that share the same sequence as the variant repeats just at the fusion point. Furthermore, the alleles that harboured the complex mutations were accompanied by simple mutations in repeats proximal to the breakpoints such additional changes have not been seen in ALT cell lines or tumours. Thus, it is likely that the unknown liposarcomas maintain their telomeres based on a recombination-like mechanism that has some differences to the ALT pathway. Recently, it was found that a patient with HNPCC developed a liposarcoma and this was associated with *hMSH2* mutations (Hirata *et al.*, 2006). In addition, in a study of 22 sporadic liposarcomas, it was seen that 14% have methylation of the *hMLH1* promoter (Seidel *et al.*, 2005). Perhaps the differences observed in the complex mutations of ALT and unknown liposarcomas are caused by differences in MMR. For example, in typical ALT the prevention of homeologous recombination between similar but not identical telomere repeat sequences could be alleviated by defects in MMR (Rizki and Lundblad, 2001).

In summary, here it is shown that some of the liposarcomas described as not activating or having an unknown TMM, in fact use an ALT-like pathway for telomere maintenance. This indicates that the classical markers used to determine ALT (*e.g.* presence of APBs) are not sufficient to determine the TMM in some cases.

Chapter 8

Final Discussion and Future Work

8.1 The effect of MMR on telomeres

In this work, the effect of *MSH2*, *MLH1* and *MSH6* genes on telomere stability was determined by analysing the distribution of telomere variant repeats in clones from cancer-derived cell lines with defects in one MMR gene (see Chapter 3). The analysis was carried out at the Xp/Yp and 12q telomeres and gave a mutation frequency of 3.6% per telomere (Table 3.4) which was significantly higher than seen in MMR proficient primary and telomerase positive cell lines ($P < 0.001$ and $P = 0.0015$ respectively). These results showed that MMR defects contribute to telomere instability in cells. In order to investigate the effect of MMR *in vivo*, two colorectal tumours with characterised mutations in the *MSH2* gene were screened for mutations at the Xp/Yp telomere. Around 100 molecules were analysed in each of the two *MSH2*-defective tumours using a single molecule approach that consisted of single telomere amplification and mutation detection using the TVR-PCR technique. The combined mutation frequency found across both tumours was 1.2% per telomere.

It was reported previously that the telomere mutation frequency in colorectal tumours with MSI (determined by analysing the stability of the polyA track in the *TGFβRII* gene) and therefore likely to contain MMR defects, was 18.6% while five colorectal tumours with characterised mutations in the *MSH2* gene had a mutation frequency of 35% per allele (Pickett *et al.*, 2004). These mutation frequencies are statistically different to the results obtained in this work (Fisher's two tailed exact test $P < 0.001$). However, in the study carried out by Pickett *et al* the mutation frequencies were determined by comparing normal colon or blood samples with colorectal cancers. Therefore mutation frequency was compared between MMR proficient and MMR deficient samples while in this work the mutation frequency was determined by clonal analysis of MMR deficient cell lines or by single molecule analysis of telomeres within tumours that were already MMR deficient. Therefore, it is possible that the higher mutation frequency reported in Pickett *et al* was the result of higher genomic instability during the transition from normal colon to carcinoma.

In order to determine whether the mutation at telomeres in cancer cell lines and tumours were produced by defects in MMR alone and are not affected by other mutations usually found in tumours (*e.g.* mutations at *TP53*, *BAX*, *APC*, *K-RAS*), downregulation of

MSH2 was carried out in a normal fibroblast cell line expressing telomerase. The results showed that indeed reduction in *MSH2* expression alone is able to generate instability at microsatellites and telomeres (Chapter 5). Furthermore, it was observed that the telomere length in two clones with *MSH2* downregulation was reduced dramatically in a period of 3 population doublings (at least 350 bp per PD) suggesting that reduced *MSH2* expression causes instability at telomeres and may also contribute to telomere shortening. It is not clear why there was such a reduction in telomere length if telomerase was active in these clones; however it has been suggested that telomerase adds *de novo* telomere repeats preferentially to short telomeres (Grobelny *et al.*, 2001). Therefore it is possible that the telomere shortening observed at the Xp/Yp telomere does not affect other telomere ends. Therefore, telomere length could be analysed at other telomere ends such as the 17p which is reported to be shorter than the median telomere length (Martens *et al.*, 1998; Britt-Compton *et al.*, 2006).

The reduction of telomere length is consistent with the finding that most of the telomere mutations in samples with MMR defects were small deletions. For example, 63% of the mutations found by Pickett *et al* (2004) and ~50% of the mutations found in this work were deletions of one or more telomere repeats compared to 17% and 19% of insertions respectively. Although telomere shortening was seen upon downregulation of *MSH2*, only 2 clones were analysed, therefore this study could be extended by generating more clones with stable downregulation of *MSH2* to confirm preliminary observations that *MSH2* downregulation leads to telomere shortening. Erosion of telomere length has been observed in colorectal cancers (with MSI) when compared to telomere length in the matched normal colon (Takagi *et al.*, 2000 and this work). Therefore, it is possible that colorectal tumours with MSI have an elevate telomere instability leading to telomere shortening. However, once the tumour is established, the genome stabilises (discussed below) and there is a reduction in telomere mutations possibly influenced by the activation of the enzyme telomerase (Figure 8.1).

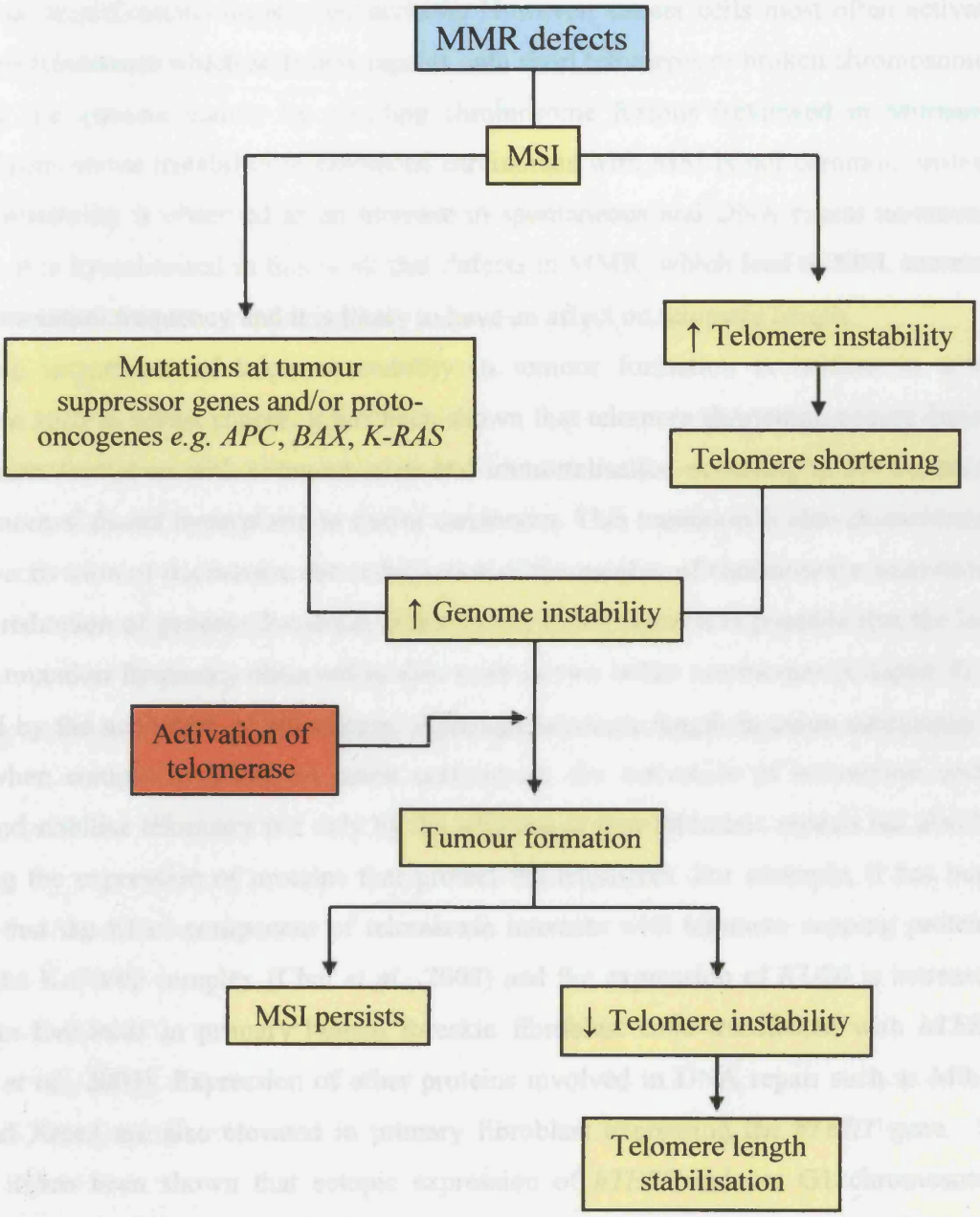


Figure 8.1 Model of telomere and microsatellite instability in cells with defects in the MMR pathway. Increase in genome and telomere instability (↑) occurs during tumourigenesis, however after tumour formation, there is a decrease (↓) in telomere instability. The activation of telomerase stabilises telomere length and may keep telomere mutations at low levels.

Genetic instability has been proposed to be a key player in the development of cancer and is observed at chromosomal level as loss or gain of whole or parts of chromosomes (Lengauer *et al.*, 1998). One mechanism of genome instability is generated by the loss of telomere repeats leading to chromosome fusions. These fusions can result in large terminal

deletions or amplifications during cell division. However, cancer cells most often activate the enzyme telomerase which adds new repeats onto short telomeres or broken chromosomes stabilising the genome mainly by avoiding chromosome fusions (reviewed in Murnane, 2006). Chromosomal instability in colorectal carcinomas with MSI is not common, instead genomic instability is observed as an increase in spontaneous and DNA repeat mutations. However, it is hypothesised in this work that defects in MMR, which lead to MSI, increase telomere mutation frequency and it is likely to have an affect on telomere length.

The importance of telomere stability in tumour formation is evident in some carcinomas such as breast cancer. It has been shown that telomere shortening occurs during breast cancer formation with telomere crisis and immortalisation occurring in the transition between normal ductal hyperplasia to ductal carcinoma. This transition is also characterised by the re-activation of telomerase, the stabilisation of the number of chromosome aberrations and thus reduction of genome instability (Chin *et al.*, 2004). Thus it is possible that the low telomere mutation frequency observed in this work in two colon carcinomas (Chapter 4) is promoted by the activation of telomerase. Although telomere length in colon carcinoma is eroded when compared to normal colon counterpart, the activation of telomerase could protect and stabilise telomeres not only by the addition of new telomeric repeats but also by promoting the expression of proteins that protect the telomeres. For example, it has been reported that the hTert component of telomerase interacts with telomere capping proteins such as the Ku70/80 complex (Chai *et al.*, 2002) and the expression of *KU80* is increased more than five-folds in primary human foreskin fibroblast cells transfected with *hTERT* (Sharma *et al.*, 2003). Expression of other proteins involved in DNA repair such as Mlh1, Msh6 and Xrcc3 are also elevated in primary fibroblast expressing the *hTERT* gene. In addition it has been shown that ectopic expression of *hTERT* induces G1 chromosome stability and enhanced DNA repair, functions of telomerase that are independent of its role of elongating telomeres (Sharma *et al.*, 2003).

The *MSH2* downregulation carried out in Chapter 5 was performed in a normal cell line that was transfected with the *hTERT* component of the telomerase (CCD34-Lu/hTERT). It would be interesting to perform similar studies but with the primary cell line (CCD34-Lu) without expressing *hTERT* since *hTERT* expression can potentially stabilise the telomeres by adding new repeats to the array or even by elevating the expression of DNA repair genes (Sharma *et al.*, 2003). However it is difficult to perform these experiments, firstly because it is hard to transfect primary cell lines and secondly because of the limited replicative potential of normal cell lines. Another possibility would be to increase the replicative

capacity of CCD34-Lu by transforming it with the simian virus 40 (SV40), however this would lead to alterations in other genes such as *TP53* and *RB*.

8.2 Effect of age on telomere stability in colon samples

In this work, the effect of age on telomere stability was assayed in some samples of colon. Telomere mutation analysis was assayed in three different normal colon samples; sample 2N derived from a 66-year old individual, sample 20N from an 81-year old individual and sample 23N from an 88-year old individual. Telomere instability was found in samples 20N and 23N however it was only seen in telomeres that contained CTAGGG repeats. Telomere instability was not found at the Xp/Yp telomere in samples 2N and at the 12q telomere in sample 23N where no CTAGGG repeats were present. Thus, it was hypothesized that the high level of instability observed in samples 20 and 23 was due to the presence of CTAGGG repeats but not related to age. Further screening will have to be carried out including more samples with a wider range of age to determine if telomere instability is affected by the age of an individual. It has been shown that microsatellite instability increases with age (Coolbaugh-Murphy *et al.*, 2005) thus; it is possible that telomere instability also increases with age; however the screening of approximately 100 molecules carried out in this work may not be enough to detect such a trend. Thus, more molecules and samples would have to be analysed to confirm the observations in this work.

8.3 Telomere instability in CTAGGG-containing telomeres

As mentioned above, two colon samples (20 and 23) with CTAGGG repeats in one Xp/Yp telomere-allele contained mutations with a frequency of 3.8% per haploid genome. Most of these mutation events consisted of insertions and deletions of a few telomere repeats involving the CTAGGG or adjacent repeats. Although two mutations were identified that involved the deletion of a larger proportion of the Xp/Yp telomere, in one case the deletion event involved more than 60 telomere repeats. These larger deletions were not observed in telomere mutation analysis of CTAGGG-containing telomeres in the germ-line, suggesting there may be differences in the mutation process. It would be interesting to screen more molecules for mutations to investigate whether there is a difference between the mutation process in the germ-line and in somatic cells. However, all the mutations found in this work and described previously (Pickett, 2002; Hills, 2004) could be explained by intra-allelic events (*e.g.* replication slippage, t-loop excisions or unequal sister-chromatid exchanges).

By analysing the telomere maps of the mutations found in the germ-line and in somatic tissue containing the CTAGGG repeats, it was found that telomeres containing blocks of CTAGGG repeats are more likely to mutate. Thus, it is possible that CTAGGG repeats form secondary structures during replication (*e.g.* G-quadruplex or hairpins). However it is not clear why only blocks of CTAGGG repeats cause such instability and other variant repeats like TCAGGG do not. It is therefore possible that secondary structures are formed in CTAGGG repeats but that CTAGGG-containing secondary structure have reduced binding affinity for proteins that resolved such structures (*e.g.* hnRPN 1 and hnRNP D, Wrn, Bloom) leading to mutations via replication fork stalling or replication slippage. Another possibility is that CTAGGG repeats have reduced binding affinity for Trf1, Trf2 and associated shelterin proteins affecting replication or repair mechanisms and thus increasing telomere instability. Therefore, it would be interesting to investigate whether $(CTAGGG)_n$ repeats have reduced binding affinity for proteins that unfold secondary structures or for Tfr1 and Trf2 proteins.

8.4 Telomere maintenance mechanism in liposarcomas

A large proportion of liposarcomas (~50%) do not use the ALT pathway or express telomerase to maintain their telomeres. In this work, telomere mutation analysis was performed in liposarcomas where no telomere maintenance mechanism (TMM) had been described. Five liposarcomas with unknown TMM were analysed for mutations at the Xp/Yp telomere and two of them were found to contain complex mutations similar to the ones observed in ALT cells (Varley *et al.*, 2002). These findings indicate that some liposarcomas described in the literature as not having activated a TMM are indeed using a recombination-like pathway to maintain their telomeres. Interestingly, these liposarcomas do not have APBs suggesting these structures are not necessary for the ALT-like process in some cancers.

The two complex mutations found in liposarcomas with unknown TMM apparently contained fusion points with sequence homology between the progenitor telomere array and the telomere repeats that were added. This possible homology between the recipient and donating repeat array has not been identified to date in cells that use the classic ALT pathway to maintain their telomeres. To extend this analysis and to elucidate whether these tumours use an ALT-like mechanism or a novel pathway it would be necessary to obtain more mutant molecules in the liposarcomas where mutations were already found or other liposarcomas with unknown TMM. It is possible that the differences observed in the complex mutations between liposarcomas using an ALT pathway and an unknown

mechanism are occasioned by differences in the levels of MMR expression, as MMR proteins are involved in preventing homeologous recombination. Thus, samples with reduced MMR might promote recombination events between telomeres that do not share similarities in their variant telomere repeat distribution. It would be interesting to quantify the expression of MMR proteins in samples that harbour complex mutations with and without homology in the repeats that are added.

It would also be interesting to perform similar analysis in other sarcomas such as osteosarcomas and Ewing's sarcomas where 18% and 30% of tumours respectively do not show any TMM mechanism (Ulaner *et al.*, 2003 and 2004). In addition, a human cell line (C3-cl6) which was derived from an ALT cell line (WI38-VA13/2RA) has been reported to maintain its telomeres in the absence of extrachromosomal DNA circles, APBs or telomerase activity (Cerone *et al.*, 2005) but this cell line has elevated telomere sister chromosome exchanges (T-SCE) as seen in ALT cells. As this cell line shows similar features as the liposarcomas with unknown TMM analysed in this work it would be interesting to use telomere mapping to investigate telomere mutation processes in this cell line.

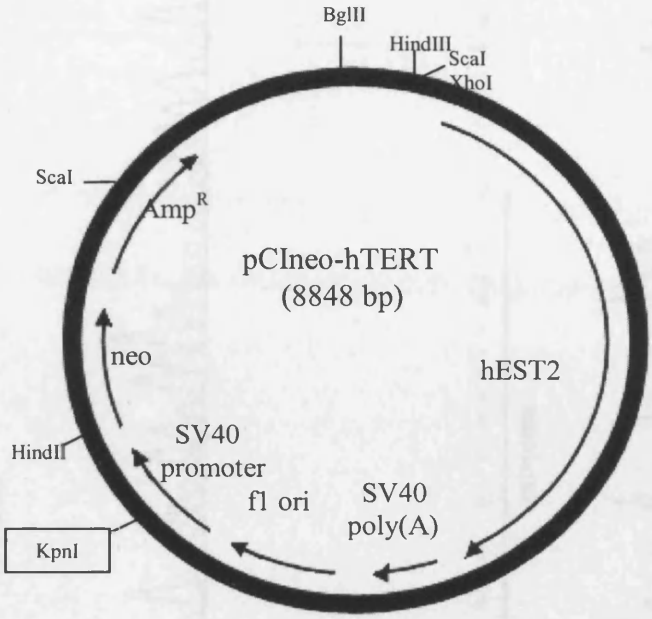
Appendix I Primer list

Primer name	Primer sequence 5′ to 3′
12q telomere	
12qA	GGGACAGCATATTCTGGTTACC
12qB	ATTTTCATTGCTGTCTTAGCACTGCAC
12qAreverse	CCCCAAAATATAATGGTAGGTAACC
pKSRV-2D	ATCCTAGCAAAGCTGAGAACTCAG
12q-197A	GGGAGATCCACACCGTAGCA
12q-197G	TGGGAGATCCACACCGTAGCG
12qnull3	GATGTCTGAGTGGATTCAGACATG
Xp/Yp telomere	
TSK8C	GCGGTACCAGGGACCGGGACAAATAGAC
TSK8G	CGGAATTCCAGACACACTAGGACCCTGA
TSK8J	GAATTCCTGGGGACTGCGGATG
TSK8M	ACCAGGTTTTCCAGTGTGTT
TSK8N	TCAGGAGTGAAGCTGCAGACC
TS-30A	CTGCTTTTATTCTCTAATCTGCTCCCA
TS-30T	CTGCTTTTATTCTCTAATCTGCTCCCT
16p/q and CB54 telomere	
Nitu14eA	GATCCAGCCAAGGCTGTCATC
Nitu14eB	GGAGACGCTCCTGACTCAGAG
Nitu14eC	GCTGTGATGCGGCTGGCACC
CB38A-F	GATCTCTCCATGCTGAGTCCAG
CB38A-R	GCTAAATTCCAGAAAGGTGTAC
Telomere repeat primers	
TAG-TelW	TCATGCGTCCATGGTCCGGACCCTTACCCTTACCCTNACCCTA
TAG-TelX	TCATGCGTCCATGGTCCGGACCCTTACCCTTACCCTNACCCTC

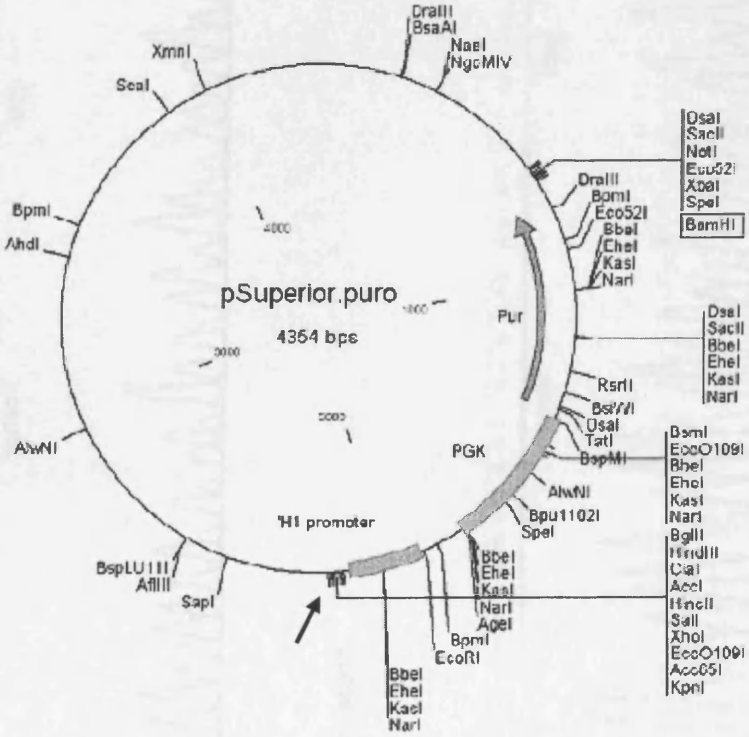
TAG-TelY	TCATGCGTCCATGGTCCGGACCCTTACCCTTACCCTNACCCTG
TAG-TelJ	TCATGCGTCCATGGTCCGGAACCCCAACCCCAACCCCAACCCC
TAG-TelCTA2	TCATGCGTCCATGGTCCGGACCCTTACCCTTRCCCTARCCCTAG
TelC	GGCCATCGATGAATTCTAACCCTAACCCTAACCCTAA
STELA primers	
Telorette	TGCTCCGTGCATGTGGCATCTAACCCT
Teltail	TGCTCCGTGCATGTGGCATC
12q-STELA	CGAAGCAGCATTCTCCTCAG
Other primers	
pSUP-1	CAGAAAGCAAGGAGCAAAG
pSUP-2	TGTGGAATTGTGAGCGGATA
PRSuper	CGAACGCTGACGTCATC
Bat26-F	TGACTACTTTTGACTTCAGCC
Bat26-R	AACCATTCAACATTTTTAACCC

Appendix II Plasmid maps

A)



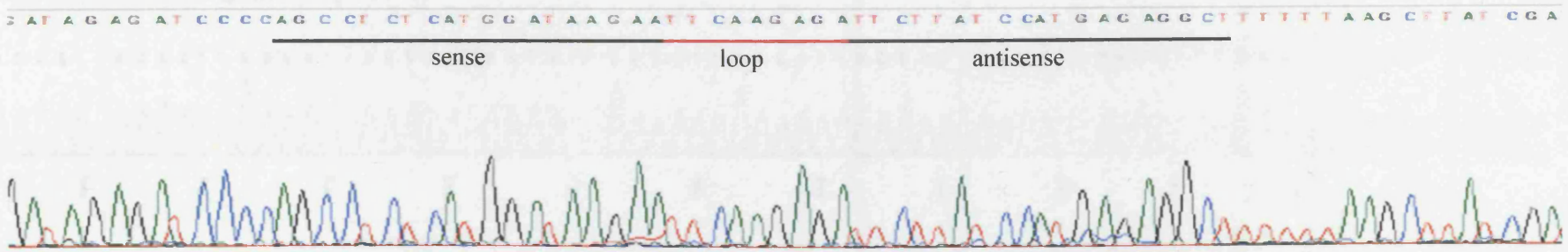
B)



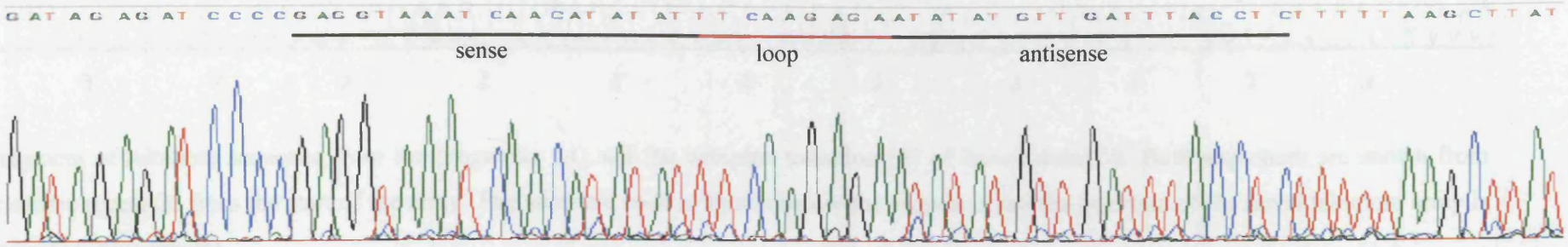
pCIneo-hTERT (A) and pSuperior.puro plasmids (B). Restriction enzymes used to linearise the plasmids before transfection into human cells are shown in boxes. The arrow in pSuperior.puro indicates the integration site for the shMSH2 oligonucleotides.

Appendix III Short-hairpin MSH2 sequences

A) shMSH2-1 sequence



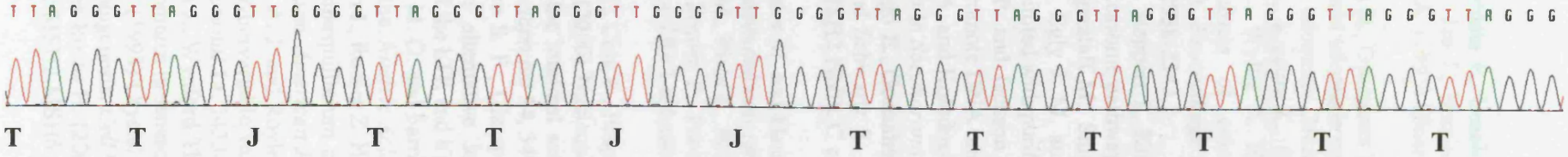
B) shMSH2-2 sequence



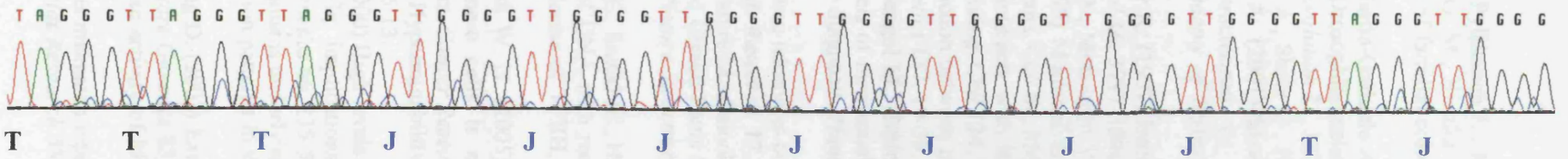
Sequencing of the shMSH2-1 (*A*) and shMSH2-2 (*B*) oligonucleotides after insertion into pSuperior.puro plasmid. The sense and antisense sequences are underline (black) as well as the short-hairpin loop (red).

Appendix IV Liposarcoma 58 sequencing

A) Liposarcoma 58 Progenitor sequence



B) Liposarcoma 58 Complex mutation sequence



Fragment of telomere sequence from the progenitor (*A*) and the complex mutation (*B*) of liposarcoma 58. Both sequences are shown from telomere repeat 25 from the start of the array. The telomere code is shown below the sequence and the insertion of the novel telomere array is indicated (blue letters).

References

Aaltonen L. A., Peltomaki P., Leach F. S., Sistonen P., Pylkkanen L., Mecklin J. P., Jarvinen H., Powell S. M., Jen J., Hamilton S. R., Petersen G. M., Kinzler K. W., Vogelstein B. and Delachapelle A. (1993) Clues to the pathogenesis of familial colorectal-cancer. *Science* **260**, 812-816.

Abad J. P., De Pablos B., Osoegawa K., De Jong P. J., Martin-Gallardo A. and Villasante A. (2004) TAHRE, a novel telomeric retrotransposon from Drosophila melanogaster, reveals the origin of Drosophila telomeres. *Molecular Biology and Evolution* **21**, 1620-1624.

Abdel-Rahman W. M., Katsura K., Rens W., Gorman P. A., Sheer D., Bicknell D., Bodmer W. F., Arends M. J., Wyllie A. H. and Edwards P. A. (2001) Spectral karyotyping suggests additional subsets of colorectal cancers characterized by pattern of chromosome rearrangement. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 2538-2543.

Aberle H., Bauer A., Stappert J., Kispert A. and Kemler R. (1997) beta-catenin is a target for the ubiquitin-proteasome pathway. *The EMBO Journal* **16**, 3797-3804.

Aboussekhra A., Biggerstaff M., Shivji M. K., Vilpo J. A., Moncollin V., Podust V. N., Protic M., Hubscher U., Egly J. M. and Wood R. D. (1995) Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell* **80**, 859-868.

Ahnesorg P., Smith P. and Jackson S. P. (2006) XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell* **124**, 301-313.

Alani E., Reenan R. A. and Kolodner R. D. (1994) Interaction between mismatch repair and genetic recombination in *Saccharomyces cerevisiae*. *Genetics* **137**, 19-39.

Alazzouzi H., Domingo E., Gonzalez S., Blanco I., Armengol M., Espin E., Plaja A., Schwartz S., Capella G. and Schwartz S., Jr. (2005) Low levels of microsatellite instability characterize MLH1 and MSH2 HNPCC carriers before tumor diagnosis. *Human Molecular Genetics* **14**, 235-239.

Allshire R. C., Dempster M. and Hastie N. D. (1989) Human telomeres contain at least 3 types of G-rich repeat distributed non-randomly. *Nucleic Acids Research.* **17**, 4611-4627.

Amiard S., Doudeau M., Pinte S., Poulet A., Lenain C., Faivre-Moskalenko C., Angelov D., Hug N., Vindigni A., Bouvet P., Paoletti J., Gilson E. and Giraud-Panis M. J. (2007) A topological mechanism for TRF2-enhanced strand invasion. *Nature Structural & Molecular Biology* **14**, 147-154.

Araujo S. J., Tirode F., Coin F., Pospiech H., Syvaoja J. E., Stucki M., Hubscher U., Egly J. M. and Wood R. D. (2000) Nucleotide excision repair of DNA with recombinant human proteins: definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes & Development* **14**, 349-359.

Atkinson S. P., Hoare S. F., Glasspool R. M. and Keith W. N. (2005) Lack of telomerase gene expression in alternative lengthening of telomere cells is associated with chromatin remodeling of the hTR and hTERT gene promoters. *Cancer Research* **65**, 7585-7590.

Bacon A. L., Dunlop M. G. and Farrington S. M. (2001) Hypermutability at a poly(A/T) tract in the human germline. *Nucleic Acids Research* **29**, 4405-13.

Baird D. M., Coleman J., Rosser Z. H. and Royle N. J. (2000) High levels of sequence polymorphism and linkage disequilibrium at the telomere of 12q: implications for telomere biology and human evolution. *American Journal of Human Genetics* **66**, 235-50.

Baird D. M., Jeffreys A. J. and Royle N. J. (1995) Mechanisms underlying telomere repeat turnover, revealed by hypervariable variant repeat distribution patterns in the human Xp/Yp telomere. *The EMBO Journal* **14**, 5433-5443.

Baird D. M., Rowson J., Wynford-Thomas D. and Kipling D. (2003) Extensive allelic variation and ultrashort telomeres in senescent human cells. *Nature Genetics* **33**, 203-7.

Ban C. and Yang W. (1998) Crystal structure and ATPase activity of MutL: implications for DNA repair and mutagenesis. *Cell* **95**, 541-552.

Banerjee S. and Flores-Rozas H. (2005) Cadmium inhibits mismatch repair by blocking the ATPase activity of the MSH2-MSH6 complex. *Nucleic Acids Research* **33**, 1410-1419.

Bartkova J., Horejsi Z., Koed K., Kramer A., Tort F., Zieger K., Guldberg P., Sehested M., Nesland J. M., Lukas C., Orntoft T., Lukas J. and Bartek J. (2005) DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis. *Nature* **434**, 864-870.

Baulcombe D. C. (1996) RNA as a target and an initiator of post-transcriptional gene silencing in transgenic plants. *Plant Molecular Biology* **32**, 79-88.

Baumann P., Benson F. E. and West S. C. (1996) Human Rad51 protein promotes ATP-dependent homologous pairing and strand transfer reactions in vitro. *Cell* **87**, 757-766.

Baumann P. and Cech T. R. (2000) Protection of telomeres by the Ku protein in fission yeast. *Molecular Biology of the Cell* **11**, 3265-3275.

Baur J. A., Zou Y., Shay J. W. and Wright W. E. (2001) Telomere position effect in human cells. *Science* **292**, 2075-7.

Bechter O. E., Zou Y., Shay J. W. and Wright W. E. (2003) Homologous recombination in human telomerase-positive and ALT cells occurs with the same frequency. *EMBO Reports* **4**, 1138-43.

Bechter O. E., Zou Y., Walker W., Wright W. E. and Shay J. W. (2004) Telomeric recombination in mismatch repair deficient human colon cancer cells after telomerase inhibition. *Cancer Research* **64**, 3444-51.

Benetti R., Garcia-Cao M. and Blasco M. A. (2007) Telomere length regulates the epigenetic status of mammalian telomeres and subtelomeres. *Nature Genetics* **39**, 243-250.

Bertuch A. A. and Lundblad V. (2003) Which end: dissecting Ku's function at telomeres and double-strand breaks. *Genes & Development* **17**, 2347-50.

Bhattacharyya N. P., Ganesh A., Phear G., Richards B., Skandalis A. and Meuth M. (1995) Molecular analysis of mutations in mutator colorectal carcinoma cell lines. *Human Molecular Genetics* **4**, 2057-2064.

Bhattacharyya N. P., Skandalis A., Ganesh A., Groden J. and Meuth M. (1994) Mutator phenotypes in human colorectal carcinoma cell lines. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 6319-6323.

Bi X., Wei S. C. and Rong Y. S. (2004) Telomere protection without a telomerase; the role of ATM and Mre11 in Drosophila telomere maintenance. *Current Biology* **14**, 1348-1353.

Bianchi A., Stansel R. M., Fairall L., Griffith J. D., Rhodes D. and de Lange T. (1999) TRF1 binds a bipartite telomeric site with extreme spatial flexibility. *The EMBO Journal* **18**, 5735-44.

Billy E., Brondani V., Zhang H., Muller U. and Filipowicz W. (2001) Specific interference with gene expression induced by long, double-stranded RNA in mouse embryonal teratocarcinoma cell lines. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 14428-14433.

Bjerknes M. and Cheng H. (1999) Clonal analysis of mouse intestinal epithelial progenitors. *Gastroenterology* **116**, 7-14.

Blackburn E. H. and Gall J. G. (1978) A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in *Tetrahymena*. *Journal of Molecular Biology* **120**, 33-53.

Blake C., Tsao J. L., Wu A. and Shibata D. (2001) Stepwise deletions of polyA sequences in mismatch repair-deficient colorectal cancers. *The American Journal of Pathology* **158**, 1867-1870.

Blasco M. A. (2007) The epigenetic regulation of mammalian telomeres. *Nature Reviews Genetics* **8**, 299-309.

Bocker T., Diermann J., Friedl W., Gebert J., Holinski-Feder E., Karner-Hanusch J., von Knebel-Doeberitz M., Koelble K., Moeslein G., Schackert H. K., Wirtz H. C., Fishel R. and Ruschoff J. (1997) Microsatellite instability analysis: a multicenter study for reliability and quality control. *Cancer Research* **57**, 4739-4743.

Bodnar A. G., Ouellette M., Frolkis M., Holt S. E., Chiu C. P., Morin G. B., Harley C. B., Shay J. W., Lichtsteiner S. and Wright W. E. (1998) Extension of life-span by introduction of telomerase into normal human cells. *Science* **279**, 349-352.

Boland C. R., Thibodeau S. N., Hamilton S. R., Sidransky D., Eshleman J. R., Burt R. W., Meltzer S. J., Rodriguez-Bigas M. A., Fodde R., Ranzani G. N. and Srivastava S. (1998) A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Research* **58**, 5248-57.

Borden E. C., Baker L. H., Bell R. S., Bramwell V., Demetri G. D., Eisenberg B. L., Fletcher C. D., Fletcher J. A., Ladanyi M., Meltzer P., O'Sullivan B., Parkinson D. R., Pisters P. W., Saxman S., Singer S., Sundaram M., van Oosterom A. T., Verweij J., Waalen J., Weiss S. W. and Brennan M. F. (2003) Soft tissue sarcomas of adults: state of the translational science. *Clinical Cancer Research* **9**, 1941-1956.

Borresen A. L., Lothe R. A., Meling G. I., Lystad S., Morrison P., Lipford J., Kane M. F., Rognum T. O. and Kolodner R. D. (1995) Somatic mutations in the hMSH2 gene in microsatellite unstable colorectal carcinomas. *Human Molecular Genetics* **4**, 2065-72.

Bowman G. D., O'Donnell M. and Kuriyan J. (2004) Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature* **429**, 724-730.

Branch P., Hampson R. and Karran P. (1995) DNA mismatch binding defects, DNA damage tolerance, and mutator phenotypes in human colorectal carcinoma cell lines. *Cancer Research* **55**, 2304-2309.

Britt-Compton B., Rowson J., Locke M., Mackenzie I., Kipling D. and Baird D. M. (2006) Structural stability and chromosome-specific telomere length is governed by cis-acting determinants in humans. *Human Molecular Genetics* **15**, 725-33.

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

Brown K. D., Rathi A., Kamath R., Beardsley D. I., Zhan Q., Mannino J. L. and Baskaran R. (2003) The mismatch repair system is required for S-phase checkpoint activation. *Nature Genetics* **33**, 80-84.

Brown P. W., Judis L., Chan E. R., Schwartz S., Seftel A., Thomas A. and Hassold T. J. (2005) Meiotic synapsis proceeds from a limited number of subtelomeric sites in the human male. *American Journal of Human Genetics* **77**, 556-66.

Bryan T. M., Marusic L., Bacchetti S., Namba M. and Reddel R. R. (1997) The telomere lengthening mechanism in telomerase-negative immortal human cells does not involve the telomerase RNA subunit. *Human Molecular Genetics* **6**, 921-926.

Buck D., Malivert L., de Chasseval R., Barraud A., Fondaneche M. C., Sanal O., Plebani A., Stephan J. L., Hufnagel M., le Deist F., Fischer A., Durandy A., de Villartay J. P. and Revy P. (2006) Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell* **124**, 287-299.

Bugreev D. V., Mazina O. M. and Mazin A. V. (2006) Rad54 protein promotes branch migration of Holliday junctions. *Nature* **442**, 590-593.

Calabrese P., Tsao J. L., Yatabe Y., Salovaara R., Mecklin J. P., Jarvinen H. J., Aaltonen L. A., Tavare S. and Shibata D. (2004) Colorectal pretumor progression before and after loss of DNA mismatch repair. *The American Journal of Pathology* **164**, 1447-1453.

Campisi J. (2005) Suppressing cancer: the importance of being senescent. *Science* **309**, 886-887.

Celli G. B. and de Lange T. (2005) DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. *Nature Cell Biology* **7**, 712-718.

Cerone M. A., Autexier C., Londono-Vallejo J. A. and Bacchetti S. (2005) A human cell line that maintains telomeres in the absence of telomerase and of key markers of ALT. *Oncogene* **24**, 7893-7901.

Cesare A. J. and Griffith J. D. (2004) Telomeric DNA in ALT cells is characterized by free telomeric circles and heterogeneous t-loops. *Molecular and Cellular Biology* **24**, 9948-9957.

Chai W., Du Q., Shay J. W. and Wright W. E. (2006a) Human telomeres have different overhang sizes at leading versus lagging strands. *Molecular Cell* **21**, 427-435.

Chai W., Ford L. P., Lenertz L., Wright W. E. and Shay J. W. (2002) Human Ku70/80 associates physically with telomerase through interaction with hTERT. *The Journal of Biological Chemistry* **277**, 47242-47247.

Chai W., Sfeir A. J., Hoshiyama H., Shay J. W. and Wright W. E. (2006b) The involvement of the Mre11/Rad50/Nbs1 complex in the generation of G-overhangs at human telomeres. *EMBO Reports* **7**, 225-230.

Chambers S. R., Hunter N., Louis E. J. and Borts R. H. (1996) The mismatch repair system reduces meiotic homeologous recombination and stimulates recombination-dependent chromosome loss. *Molecular and Cellular Biology* **16**, 6110-6120.

Chan T. L., Yuen S. T., Kong C. K., Chan Y. W., Chan A. S., Ng W. F., Tsui W. Y., Lo M. W., Tam W. Y., Li V. S. and Leung S. Y. (2006) Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. *Nature Genetics* **38**, 1178-1183.

Chen J. and Sadowski I. (2005) Identification of the mismatch repair genes PMS2 and MLH1 as p53 target genes by using serial analysis of binding elements. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 4813-4818.

Chen J. L., Blasco M. A. and Greider C. W. (2000) Secondary structure of vertebrate telomerase RNA. *Cell* **100**, 503-514.

Chen J. L. and Greider C. W. (2004) An emerging consensus for telomerase RNA structure. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14683-14684.

Chen Q., Ijpma A. and Greider C. W. (2001) Two survivor pathways that allow growth in the absence of telomerase are generated by distinct telomere recombination events. *Molecular and Cellular Biology* **21**, 1819-27.

Chen W. and Jinks-Robertson S. (1999) The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. *Genetics* **151**, 1299-313.

Chiang Y. J., Nguyen M. L., Gurunathan S., Kaminker P., Tessarollo L., Campisi J. and Hodes R. J. (2006) Generation and characterization of telomere length maintenance in tankyrase 2-deficient mice. *Molecular and Cellular Biology* **26**, 2037-2043.

Chin K., de Solorzano C. O., Knowles D., Jones A., Chou W., Rodriguez E. G., Kuo W. L., Ljung B. M., Chew K., Myambo K., Miranda M., Krig S., Garbe J., Stampfer M., Yaswen P., Gray J. W. and Lockett S. J. (2004) In situ analyses of genome instability in breast cancer. *Nature Genetics* **36**, 984-8.

Chong S. S., McCall A. E., Cota J., Subramony S. H., Orr H. T., Hughes M. R. and Zoghbi H. Y. (1995) Gametic and somatic tissue-specific heterogeneity of the expanded SCA1 CAG repeat in spinocerebellar ataxia type 1. *Nature Genetics* **10**, 344-350.

Chu G., Hayakawa H. and Berg P. (1987) Electroporation for the efficient transfection of mammalian cells with DNA. *Nucleic Acids Research* **15**, 1311-1326.

Coindre J. M., Terrier P., Guillou L., Le Doussal V., Collin F., Ranchere D., Sastre X., Vilain M. O., Bonichon F. and N'Guyen Bui B. (2001) Predictive value of grade for metastasis development in the main histologic types of adult soft tissue sarcomas: a study of 1240 patients from the French Federation of Cancer Centers Sarcoma Group. *Cancer* **91**, 1914-1926.

Coleman J., Baird D. M. and Royle N. J. (1999) The plasticity of human telomeres demonstrated by a hypervariable telomere repeat array that is located on some copies of 16p and 16q. *Human Molecular Genetics* **8**, 1637-1646.

Coleman M. G., Gough A. C., Bunyan D. J., Braham D., Eccles D. M. and Primrose J. N. (2001) Minisatellite instability is found in colorectal tumours with mismatch repair deficiency. *British Journal of Cancer* **85**, 1486-91.

Colgin L. M., Wilkinson C., Englezou A., Kilian A., Robinson M. O. and Reddel R. R. (2000) The hTERTalpha splice variant is a dominant negative inhibitor of telomerase activity. *Neoplasia* **2**, 426-32.

Constantin N., Dzantiev L., Kadyrov F. A. and Modrich P. (2005) Human mismatch repair: reconstitution of a nick-directed bidirectional reaction. *The Journal of Biological Chemistry* **280**, 39752-39761.

Constantinou A., Tarsounas M., Karow J. K., Brosh R. M., Bohr V. A., Hickson I. D. and West S. C. (2000) Werner's syndrome protein (WRN) migrates Holliday junctions and co-localizes with RPA upon replication arrest. *EMBO reports* **1**, 80-84.

Cook B. D., Dynek J. N., Chang W., Shostak G. and Smith S. (2002) Role for the related poly(ADP-Ribose) polymerases tankyrase 1 and 2 at human telomeres. *Molecular and Cellular Biology* **22**, 332-342.

Coolbaugh-Murphy M., Maleki A., Ramagli L., Frazier M., Lichtiger B., Monckton D. G., Siciliano M. J. and Brown B. W. (2004) Estimating mutant microsatellite allele frequencies in somatic cells by small-pool PCR. *Genomics* **84**, 419-430.

Coolbaugh-Murphy M. I., Xu J., Ramagli L. S., Brown B. W. and Siciliano M. J. (2005) Microsatellite instability (MSI) increases with age in normal somatic cells. *Mechanisms of Ageing and Development* **126**, 1051-1059.

Cooper D. L., Lahue R. S. and Modrich P. (1993) Methyl-directed mismatch repair is bidirectional. *The Journal of Biological Chemistry* **268**, 11823-11829.

Costa A., Daidone M. G., Daprai L., Villa R., Cantu S., Pilotti S., Mariani L., Gronchi A., Henson J. D., Reddel R. R. and Zaffaroni N. (2006) Telomere maintenance mechanisms in liposarcomas: association with histologic subtypes and disease progression. *Cancer Research* **66**, 8918-24.

Cottu P. H., Muzeau F., Estreicher A., Flejou J. F., Iggo R., Thomas G. and Hamelin R. (1996) Inverse correlation between RER+ status and p53 mutation in colorectal cancer cell lines. *Oncogene* **13**, 2727-2730.

Counter C. M., Avilion A. A., Lefeuvre C. E., Stewart N. G., Greider C. W., Harley C. B. and Bacchetti S. (1992) Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *The EMBO Journal* **11**, 1921-1929.

Crabbe L., Jauch A., Naeger C. M., Holtgreve-Grez H. and Karlseder J. (2007) Telomere dysfunction as a cause of genomic instability in Werner syndrome. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 2205-2210.

Crawford D. C., Wilson B. and Sherman S. L. (2000) Factors involved in the initial mutation of the fragile X CGG repeat as determined by sperm small pool PCR. *Human Molecular Genetics* **9**, 2909-18.

Crosnier C., Stamataki D. and Lewis J. (2006) Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nature Reviews Genetics* **7**, 349-359.

d'Adda di Fagagna F., Reaper P. M., Clay-Farrace L., Fiegler H., Carr P., Von Zglinicki T., Saretzki G., Carter N. P. and Jackson S. P. (2003) A DNA damage checkpoint response in telomere-initiated senescence. *Nature* **426**, 194-8.

Dao V. and Modrich P. (1998) Mismatch-, MutS-, MutL-, and helicase II-dependent unwinding from the single-strand break of an incised heteroduplex. *The Journal of Biological Chemistry* **273**, 9202-9207.

Datta A., Hendrix M., Lipsitch M. and Jinks-Robertson S. (1997) Dual roles for DNA sequence identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 9757-9762.

Davies A. A., Masson J. Y., McIlwraith M. J., Stasiak A. Z., Stasiak A., Venkitaraman A. R. and West S. C. (2001) Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Molecular Cell* **7**, 273-282.

Day S. J., Nelson M., Rosenthal H., Vergara G. G. and Bridge J. A. (1997) Der(16)t(1;16)(q21;q13) as a secondary structural aberration in yet a third sarcoma, extraskeletal myxoid chondrosarcoma. *Genes, Chromosomes & Cancer* **20**, 425-427.

de Lange T. (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes & Development* **19**, 2100-2110.

de Wind N., Dekker M., Claij N., Jansen L., van Klink Y., Radman M., Riggins G., van der Valk M., van't Wout K. and te Riele H. (1999) HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. *Nature Genetics* **23**, 359-362.

Dikovskaya D., Schiffmann D., Newton I. P., Oakley A., Kroboth K., Sansom O., Jamieson T. J., Meniel V., Clarke A. and Nathke I. S. (2007) Loss of APC induces polyploidy as a result of a combination of defects in mitosis and apoptosis. *The Journal of Cell Biology* **176**, 183-195.

Dimri G. P., Lee X. H., Basile G., Acosta M., Scott C., Roskelley C., Medrano E. E., Linskens M., Rubelj I., Pereira-Smith O., Peacocke M. and Campisi J. (1995) A biomarker that identifies senescent human cells in culture and in aging skin *in vivo*. *Proceedings of the National Academy of Sciences of the United States of Americal* **92**, 9363-9367.

Ding H., Schertzer M., Wu X., Gertsenstein M., Selig S., Kammori M., Pourvali R., Poon S., Vulto I., Chavez E., Tam P. P., Nagy A. and Lansdorp P. M. (2004) Regulation of murine telomere length by Rtel: an essential gene encoding a helicase-like protein. *Cell* **117**, 873-86.

Dizdaroglu M. (2005) Base-excision repair of oxidative DNA damage by DNA glycosylases. *Mutation Research* **591**, 45-59.

Dokal I. (2000) Dyskeratosis congenita in all its forms. *British Journal of Haematology* **110**, 768-779.

Duckett D. R., Drummond J. T., Murchie A. I., Reardon J. T., Sancar A., Lilley D. M. and Modrich P. (1996) Human MutSalpha recognizes damaged DNA base pairs containing O6-methylguanine, O4-methylthymine, or the cisplatin-d(GpG) adduct. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 6443-6447.

Dunham M. A., Neumann A. A., Fasching C. L. and Reddel R. R. (2000) Telomere maintenance by recombination in human cells. *Nature Genetics* **26**, 447-50.

Duquette M. L., Handa P., Vincent J. A., Taylor A. F. and Maizels N. (2004) Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes & Development* **18**, 1618-1629.

Dzantiev L., Constantin N., Genschel J., Iyer R. R., Burgers P. M. and Modrich P. (2004) A defined human system that supports bidirectional mismatch-provoked excision. *Molecular Cell* **15**, 31-41.

Efstathiou J. A., Liu D., Wheeler J. M., Kim H. C., Beck N. E., Ilyas M., Karayiannakis A. J., Mortensen N. J., Kmiot W., Playford R. J., Pignatelli M. and Bodmer W. F. (1999) Mutated epithelial cadherin is associated with increased tumorigenicity and loss of adhesion and of responsiveness to the motogenic trefoil factor 2 in colon carcinoma cells. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 2316-2321.

Ejima Y., Yang L. and Sasaki M. S. (2000) Aberrant splicing of the ATM gene associated with shortening of the intronic mononucleotide tract in human colon tumor cell lines: a novel mutation target of microsatellite instability. *International Journal of Cancer* **86**, 262-268.

Elbashir S. M., Harborth J., Lendeckel W., Yalcin A., Weber K. and Tuschl T. (2001a) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-498.

Elbashir S. M., Harborth J., Weber K. and Tuschl T. (2002) Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods* **26**, 199-213.

Elbashir S. M., Lendeckel W. and Tuschl T. (2001b) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Development* **15**, 188-200.

Enokizono Y., Konishi Y., Nagata K., Ouhashi K., Uesugi S., Ishikawa F. and Katahira M. (2005) Structure of hnRNP D complexed with single-stranded telomere DNA and unfolding of the quadruplex by heterogeneous nuclear ribonucleoprotein D. *The Journal of Biological Chemistry* **280**, 18862-18870.

Eshleman J. R., Casey G., Kochera M. E., Sedwick W. D., Swinler S. E., Veigl M. L., Willson J. K. V., Schwartz S. and Markowitz S. D. (1998) Chromosome number and structure both are markedly stable in RER colorectal cancers and are not destabilized by mutation of p53. *Oncogene* **17**, 719-725.

Fasching C. L., Bower K. and Reddel R. R. (2005) Telomerase-independent telomere length maintenance in the absence of alternative lengthening of telomeres-associated promyelocytic leukemia bodies. *Cancer Research* **65**, 2722-9.

Feinberg A. P. and Vogelstein B. (1984) "A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity". *Analytical Biochemistry* **137**, 266-7.

Fire A., Xu S., Montgomery M. K., Kostas S. A., Driver S. E. and Mello C. C. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **391**, 806-811.

Fleisher A. S., Esteller M., Tamura G., Rashid A., Stine O. C., Yin J., Zou T. T., Abraham J. M., Kong D., Nishizuka S., James S. P., Wilson K. T., Herman J. G. and Meltzer S. J. (2001) Hypermethylation of the hMLH1 gene promoter is associated with microsatellite instability in early human gastric neoplasia. *Oncogene* **20**, 329-335.

Flores I., Evan G. and Blasco M. A. (2006) Genetic analysis of myc and telomerase interactions in vivo. *Molecular and Cellular Biology* **26**, 6130-6138.

Fukagawa T., Nakamura Y., Okumura K., Nogami M., Ando A., Inoko H., Saitou N. and Ikemura T. (1996) Human pseudoautosomal boundary-like sequences - expression and involvement in evolutionary formation of the present-day pseudoautosomal boundary of human sex-chromosomes. *Human Molecular Genetics* **5**, 23-32.

Garcia-Cao M., O'Sullivan R., Peters A. H., Jenuwein T. and Blasco M. A. (2004) Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nature Genetics* **36**, 94-9.

Gayet J., Zhou X. P., Duval A., Rolland S., Hoang J. M., Cottu P. and Hamelin R. (2001) Extensive characterization of genetic alterations in a series of human colorectal cancer cell lines. *Oncogene* **20**, 5025-32.

Genschel J., Littman S. J., Drummond J. T. and Modrich P. (1998) Isolation of MutSbeta from human cells and comparison of the mismatch repair specificities of MutSbeta and MutSalpha. *The Journal of Biological Chemistry* **273**, 19895-19901.

Gitlin L. and Andino R. (2003) Nucleic acid-based immune system: the antiviral potential of mammalian RNA silencing. *Journal of Virology* **77**, 7159-7165.

Glaab W. E. and Tindall K. R. (1997) Mutation rate at the hprt locus in human cancer cell lines with specific mismatch repair-gene defects. *Carcinogenesis* **18**, 1-8.

Gonzalez-Suarez E., Samper E., Flores J. M. and Blasco M. A. (2000) Telomerase-deficient mice with short telomeres are resistant to skin tumorigenesis. *Nature Genetics* **26**, 114-7.

Gonzalo S., Jaco I., Fraga M. F., Chen T., Li E., Esteller M. and Blasco M. A. (2006) DNA methyltransferases control telomere length and telomere recombination in mammalian cells. *Nature Cell Biology* **8**, 416-24.

Gottschling D. E., Aparicio O. M., Billington B. L. and Zakian V. A. (1990) Position effect at *Saccharomyces cerevisiae* telomeres: reversible repression of pol II transcription. *Cell* **63**, 751-762.

Gray J. T., Celander D. W., Price C. M. and Cech T. R. (1991) Cloning and expression of genes for the oxytricha telomere-binding protein - specific subunit interactions in the telomeric complex. *Cell* **67**, 807-814.

Gregorieff A., Pinto D., Begthel H., Destree O., Kielman M. and Clevers H. (2005) Expression pattern of Wnt signaling components in the adult intestine. *Gastroenterology* **129**, 626-638.

Greider C. W. (1990) Telomeres, telomerase and senescence. *Bioessays* **12**, 363-9.

Greider C. W. and Blackburn E. H. (1985) Identification of a specific telomere terminal transferase activity in *Tetrahymena* extracts. *Cell* **43**, 405-13.

Griffith J. D., Comeau L., Rosenfield S., Stansel R. M., Bianchi A., Moss H. and de Lange T. (1999) Mammalian telomeres end in a large duplex loop. *Cell* **97**, 503-14.

Grobelny J. V., Kulp-McEliece M. and Broccoli D. (2001) Effects of reconstitution of telomerase activity on telomere maintenance by the alternative lengthening of telomeres (ALT) pathway. *Human Molecular Genetics* **10**, 1953-61.

Groden J., Thliveris A., Samowitz W., Carlson M., Gelbert L., Albertsen H., Joslyn G., Stevens J., Spirio L. and Robertson M. (1991) Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* **66**, 589-600.

Guo S., Zhang Y., Yuan F., Gao Y., Gu L., Wong I. and Li G. M. (2006) Regulation of replication protein A functions in DNA mismatch repair by phosphorylation. *The Journal of Biological Chemistry* **281**, 21607-21616.

Hake S. (2003) MicroRNAs: a role in plant development. *Current Biology* **13**, R851-2.

Hammond S. M., Bernstein E., Beach D. and Hannon G. J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature* **404**, 293-296.

Hanaoka S., Nagadoi A. and Nishimura Y. (2005) Comparison between TRF2 and TRF1 of their telomeric DNA-bound structures and DNA-binding activities. *Protein Science* **14**, 119-130.

Harfe B. D. and Jinks-Robertson S. (2000) DNA mismatch repair and genetic instability. *Annual Review of Genetics* **34**, 359-399.

Harfe B. D., Minesinger B. K. and Jinks-Robertson S. (2000) Discrete in vivo roles for the MutL homologs Mlh2p and Mlh3p in the removal of frameshift intermediates in budding yeast. *Current Biology* **10**, 145-148.

Harley C. B., Futcher A. B. and Greider C. W. (1990) Telomeres shorten during ageing of human fibroblasts. *Nature* **345**, 458-460.

Hastie N. D., Dempster M., Dunlop M. G., Thompson A. M., Green D. K. and Allshire R. C. (1990) Telomere reduction in human colorectal carcinoma and with ageing. *Nature* **346**, 866-868.

Helleday T., Lo J., van Gent D. C. and Engelward B. P. (2007) DNA double-strand break repair: From mechanistic understanding to cancer treatment. *DNA Repair* **6**, 923-35.

Hendrickson E. A. (1997) Cell-cycle regulation of mammalian DNA double-strand-break repair. *American Journal of Human Genetics* **61**, 795-800.

Hendriks Y. M., Jagmohan-Changur S., van der Klift H. M., Morreau H., van Puijenbroek M., Tops C., van Os T., Wagner A., Ausems M. G., Gomez E., Breuning M. H., Brocker-Vriends A. H., Vasen H. F. and Wijnen J. T. (2006) Heterozygous mutations in PMS2 cause hereditary nonpolyposis colorectal carcinoma (Lynch syndrome). *Gastroenterology* **130**, 312-322.

Henricks W. H., Chu Y. C., Goldblum J. R. and Weiss S. W. (1997) Dedifferentiated liposarcoma: a clinicopathological analysis of 155 cases with a proposal for an expanded definition of dedifferentiation. *The American Journal of Surgical Pathology* **21**, 271-281.

Henson J. D., Hannay J. A., McCarthy S. W., Royds J. A., Yeager T. R., Robinson R. A., Wharton S. B., Jellinek D. A., Arbuckle S. M., Yoo J., Robinson B. G., Learoyd D. L., Stalley P. D., Bonar S. F., Yu D., Pollock R. E. and Reddel R. R. (2005) A robust assay for alternative lengthening of telomeres in tumors shows the significance of alternative lengthening of telomeres in sarcomas and astrocytomas. *Clinical Cancer Research* **11**, 217-25.

Henson J. D., Neumann A. A., Yeager T. R. and Reddel R. R. (2002) Alternative lengthening of telomeres in mammalian cells. *Oncogene* **21**, 598-610.

Herbig U., Ferreira M., Condel L., Carey D. and Sedivy J. M. (2006) Cellular senescence in aging primates. *Science* **311**, 1257.

Heyer J., Yang K., Lipkin M., Edelmann W. and Kucherlapati R. (1999) Mouse models for colorectal cancer. *Oncogene* 18, 5325-5333.

Hienonen T., Sammalkorpi H., Enholm S., Alhopuro P., Barber T. D., Lehtonen R., Nupponen N. N., Lehtonen H., Salovaara R., Mecklin J. P., Jarvinen H., Koistinen R., Arango D., Launonen V., Vogelstein B., Karhu A. and Aaltonen L. A. (2005) Mutations in two short noncoding mononucleotide repeats in most microsatellite-unstable colorectal cancers. *Cancer Research* **65**, 4607-4613.

Hills M. (2004) The Xp/Yp telomere and its adjacent DNA; structure, sequence organisation and mutation. *PhD thesis*. University of Leicester.

Hiraoka Y., Henderson E. and Blackburn E. H. (1998) Not so peculiar: fission yeast telomere repeats. *Trends in Biochemical Sciences* **23**, 126.

Hirata K., Kanemitsu S., Nakayama Y., Nagata N., Itoh H., Ohnishi H., Ishikawa H. and Furukawa Y. (2006) A novel germline mutation of MSH2 in a hereditary nonpolyposis colorectal cancer patient with liposarcoma. *The American Journal of Gastroenterology* **101**, 193-196.

Hoang J. M., Cottu P. H., Thuille B., Salmon R. J., Thomas G. and Hamelin R. (1997) BAT-26, an indicator of the replication error phenotype in colorectal cancers and cell lines. *Cancer Research* **57**, 300-303.

Hollingsworth N. M., Ponte L. and Halsey C. (1995) MSH5, a novel MutS homolog, facilitates meiotic reciprocal recombination between homologs in *Saccharomyces cerevisiae* but not mismatch repair. *Genes & Development* **9**, 1728-39.

Hsiao S. J., Poitras M. F., Cook B. D., Liu Y. and Smith S. (2006) Tankyrase 2 poly(ADP-ribose) polymerase domain-deleted mice exhibit growth defects but have normal telomere length and capping. *Molecular and Cellular Biology* **26**, 2044-2054.

Huffman K. E., Levene S. D., Tesmer V. M., Shay J. W. and Wright W. E. (2000) Telomere shortening is proportional to the size of the G-rich telomeric 3-overhang. *Journal of Biological Chemistry* **275**, 19719-19722.

Hunter N. and Borts R. H. (1997) Mlh1 is unique among mismatch repair proteins in its ability to promote crossing-over during meiosis. *Genes & Development* **11**, 1573-1582.

Ionov Y., Peinado M. A., Malkhosyan S., Shibata D. and Perucho M. (1993) Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* **363**, 558-561.

Jacob N. K., Skopp R. and Price C. M. (2001) G-overhang dynamics at Tetrahymena telomeres. *The EMBO Journal* **20**, 4299-4308.

Jagmohan-Changur S., Poikonen T., Vilkki S., Launonen V., Wikman F., Orntoft T. F., Moller P., Vasen H., Tops C., Kolodner R. D., Mecklin J. P., Jarvinen H., Bevan S., Houlston R. S., Aaltonen L. A., Fodde R., Wijnen J. and Karhu A. (2003) EXO1 variants occur commonly in normal population: evidence against a role in hereditary nonpolyposis colorectal cancer. *Cancer Research* **63**, 154-158.

Jeffreys A. J. and Neumann R. (1997) Somatic mutation processes at a human minisatellite. *Human Molecular Genetics* **6**, 129-136.

Jeffreys A. J., Tamaki K., Macleod A., Monckton D. G., Neil D. L. and Armour J. (1994) Complex gene conversion events in germline mutation at human minisatellites. *Nature Genetics* **6**, 136-145.

Jeffreys A. J., Wilson V., Neumann R. and Keyte J. (1988) Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. *Nucleic Acids Research* **16**, 10953-71.

Jeyapalan J. N., Varley H., Foxon J. L., Pollock R. E., Jeffreys A. J., Henson J. D., Reddel R. R. and Royle N. J. (2005) Activation of the ALT pathway for telomere maintenance can affect other sequences in the human genome. *Human Molecular Genetics* **14**, 1785-1794.

Jiang W. Q., Zhong Z. H., Henson J. D., Neumann A. A., Chang A. C. and Reddel R. R. (2005) Suppression of Alternative Lengthening of Telomeres by Sp100-Mediated Sequestration of the MRE11/RAD50/NBS1 Complex. *Molecular and Cellular Biology* **25**, 2708-21.

Jin Y. H., Clark A. B., Slebos R. J., Al-Refai H., Taylor J. A., Kunkel T. A., Resnick M. A. and Gordenin D. A. (2003) Cadmium is a mutagen that acts by inhibiting mismatch repair. *Nature Genetics* **34**, 326-329.

Johnson J. E. and Broccoli D. (2007) Telomere maintenance in sarcomas. *Current Opinion in Oncology* **19**, 377-382.

Johnson J. E., Varkonyi R. J., Schwalm J., Cragle R., Klein-Szanto A., Patchefsky A., Cukierman E., von Mehren M. and Broccoli D. (2005) Multiple mechanisms of telomere maintenance exist in liposarcomas. *Clinical Cancer Research* **11**, 5347-5355.

Kane M. F., Loda M., Gaida G. M., Lipman J., Mishra R., Goldman H., Jessup J. M. and Kolodner R. (1997) Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumours and mismatch repair-defective human tumour cell lines. *Cancer Research* **57**, 808-811.

Karlseder J., Broccoli D., Dai Y. M., Hardy S. and deLange T. (1999) p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science* **283**, 1321-1325.

Karlseder J., Hoke K., Mirzoeva O. K., Bakkenist C., Kastan M. B., Petrini J. H. and de Lange T. (2004) The telomeric protein TRF2 binds the ATM kinase and can inhibit the ATM-dependent DNA damage response. *PLoS Biology* **2**, E240.

Karow J. K., Constantinou A., Li J. L., West S. C. and Hickson I. D. (2000) The Bloom's syndrome gene product promotes branch migration of holliday junctions. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6504-6508.

Karran P. (2001) Mechanisms of tolerance to DNA damaging therapeutic drugs. *Carcinogenesis* **22**, 1931-1937.

Kelleher C., Kurth I. and Lingner J. (2005) Human protection of telomeres 1 (POT1) is a negative regulator of telomerase activity in vitro. *Molecular and Cellular Biology* **25**, 808-818.

Khorasanizadeh S. (2004) The nucleosome: from genomic organization to genomic regulation. *Cell* **116**, 259-272.

Kim N. W., Piatyszek M. A., Prowse K. R., Harley C. B., West M. D., Ho P. L., Coviello G. M., Wright W. E., Weinrich S. L. and Shay J. W. (1994) Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-2015.

Kirk K. E. and Blackburn E. H. (1995) An unusual sequence arrangement in the telomeres of the germ-line micronucleus in Tetrahymena thermophila. *Genes & Development* **9**, 59-71.

Kirkland S. C. (1988) Clonal origin of columnar, mucous, and endocrine cell lineages in human colorectal epithelium. *Cancer* **61**, 1359-1363.

Klobutcher L. A., Swanton M. T., Donini P. and Prescott D. M. (1981) All gene-sized DNA molecules in four species of hypotrichs have the same terminal sequence and an unusual 3' terminus. *Proceedings of the National Academy of Sciences of the United States of America* **78**, 3015-3019.

Knight J. C., Renwick P. J., Cin P. D., Van den Berghe H. and Fletcher C. D. (1995) Translocation t(12;16)(q13;p11) in myxoid liposarcoma and round cell liposarcoma: molecular and cytogenetic analysis. *Cancer Research* **55**, 24-27.

Knutson J. C. and Yee D. (1987) Electroporation: parameters affecting transfer of DNA into mammalian cells. *Analytical Biochemistry* **164**, 44-52.

Koering C. E., Pollice A., Zibella M. P., Bauwens S., Puisieux A., Brunori M., Brun C., Martins L., Sabatier L., Pulitzer J. F. and Gilson E. (2002) Human telomeric position effect is determined by chromosomal context and telomeric chromatin integrity. *EMBO Reports* **3**, 1055-1061.

Kolas N. K., Svetlanov A., Lenzi M. L., Macaluso F. P., Lipkin S. M., Liskay R. M., Greally J., Edelmann W. and Cohen P. E. (2005) Localization of MMR proteins on meiotic chromosomes in mice indicates distinct functions during prophase I. *The Journal of Cell Biology* **171**, 447-458.

Ku J. L., Yoon K. A., Kim D. Y. and Park J. G. (1999) Mutations in hMSH6 alone are not sufficient to cause the microsatellite instability in colorectal cancer cell lines. *European Journal of Cancer* **35**, 1724-1729.

Kunkel T. A. and Erie D. A. (2005) DNA mismatch repair. *Annual Review of Biochemistry* **74**, 681-710.

Laine J. P. and Egly J. M. (2006) Initiation of DNA repair mediated by a stalled RNA polymerase IIO. *The The EMBO Journal* **25**, 387-397.

Lamers M. H., Perrakis A., Enzlin J. H., Winterwerp H. H., de Wind N. and Sixma T. K. (2000) The crystal structure of DNA mismatch repair protein MutS binding to a G x T mismatch. *Nature* **407**, 711-717.

Larrivee M., LeBel C. and Wellinger R. J. (2004) The generation of proper constitutive G-tails on yeast telomeres is dependent on the MRX complex. *Genes & Development* **18**, 1391-1396.

Le Y. and Dobson M. J. (1997) Stabilization of yeast artificial chromosome clones in a rad54-3 recombination-deficient host strain. *Nucleic Acids Research* **25**, 1248-1253.

Lebbink J. H., Georgijevic D., Natrajan G., Fish A., Winterwerp H. H., Sixma T. K. and de Wind N. (2006) Dual role of MutS glutamate 38 in DNA mismatch discrimination and in the authorization of repair. *The EMBO Journal* **25**, 409-419.

Lebeau J., Fouchet P., Ory K. and Chevillard S. (2002) Down-regulation of telomerase activity after progesterone treatment of human breast cancer cells: essential role of the cell cycle status. *Anticancer Research* **22**, 2161-2166.

Lee R. C., Feinbaum R. L. and Ambros V. (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75**, 843-854.

Lee S. D. and Alani E. (2006) Analysis of interactions between mismatch repair initiation factors and the replication processivity factor PCNA. *Journal of Molecular Biology* **355**, 175-184.

Lee Y., Ahn C., Han J., Choi H., Kim J., Yim J., Lee J., Provost P., Radmark O., Kim S. and Kim V. N. (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415-419.

Lee Y., Jeon K., Lee J. T., Kim S. and Kim V. N. (2002) MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO Journal* **21**, 4663-4670.

Lee Y. S., Nakahara K., Pham J. W., Kim K., He Z., Sontheimer E. J. and Carthew R. W. (2004) Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117**, 69-81.

Lendvay T. S., Morris D. K., Sah J., Balasubramanian B. and Lundblad V. (1996) Senescence mutants of *Saccharomyces cerevisiae* with a defect in telomere replication identify three additional EST genes. *Genetics* **144**, 1399-1412.

Li B. and Comai L. (2001) Requirements for the nucleolytic processing of DNA ends by the Werner syndrome protein-Ku70/80 complex. *The Journal of Biological Chemistry* **276**, 9896-9902.

Lipkin S. M., Wang V., Jacoby R., Banerjee-Basu S., Baxevanis A. D., Lynch H. T., Elliott R. M. and Collins F. S. (2000) MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. *Nature Genetics* **24**, 27-35.

Liu Y., Masson J. Y., Shah R., O'Regan P. and West S. C. (2004) RAD51C is required for Holliday junction processing in mammalian cells. *Science* **303**, 243-6.

Loayza D. and De Lange T. (2003) POT1 as a terminal transducer of TRF1 telomere length control. *Nature* **424**, 1013-8.

Loayza D., Parsons H., Donigian J., Hoke K. and de Lange T. (2004) DNA binding features of human POT1: a nonamer 5'-TAGGGTTAG-3' minimal binding site, sequence specificity, and internal binding to multimeric sites. *The Journal of Biological Chemistry* **279**, 13241-8.

Londono-Vallejo J. A., Der-Sarkissian H., Cazes L., Bacchetti S. and Reddel R. R. (2004) Alternative lengthening of telomeres is characterized by high rates of telomeric exchange. *Cancer Research* **64**, 2324-7.

Louis E. J. (1995) The chromosome ends of *Saccharomyces cerevisiae*. *Yeast* **11**, 1553-1573.

Lund E., Guttinger S., Calado A., Dahlberg J. E. and Kutay U. (2004) Nuclear export of microRNA precursors. *Science* **303**, 95-98.

Lundblad V. (2002) Telomere maintenance without telomerase. *Oncogene* **21**, 522-31.

Lundblad V. and Blackburn E. H. (1993) An alternative pathway for yeast telomere maintenance rescues *est1-* senescence. *Cell* **73**, 347-360.

Lundin C., Erixon K., Arnaudeau C., Schultz N., Jenssen D., Meuth M. and Helleday T. (2002) Different roles for nonhomologous end joining and homologous recombination following replication arrest in mammalian cells. *Molecular and Cellular Biology* **22**, 5869-5878.

Luo K., Vega-Palas M. A. and Grunstein M. (2002) Rap1-Sir4 binding independent of other Sir, yKu, or histone interactions initiates the assembly of telomeric heterochromatin in yeast. *Genes & Development* **16**, 1528-1539.

Lynch H. T. and de la Chapelle A. (2003) Hereditary colorectal cancer. *The New England Journal of Medicine* **348**, 919-932.

Malkhosyan S., Rampino N., Yamamoto H. and Perucho M. (1996) Frameshift mutator mutations. *Nature* **382**, 499-500.

Mancuso T., Aguilar F., Pescarolo M. P., Clerico L., Russo P. and Parodi S. (1997) Mutation frequencies at codon 248 of the p53 tumour suppressor gene are not increased in colon cancer cell lines with the RER+ phenotype. *Nucleic Acids Research* **25**, 3643-3648.

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

Marcand S., Brevet V., Mann C. and Gilson E. (2000) Cell cycle restriction of telomere elongation. *Current Biology* **10**, 487-490.

Marciniak R. A., Cavazos D., Montellano R., Chen Q., Guarente L. and Johnson F. B. (2005) A novel telomere structure in a human alternative lengthening of telomeres cell line. *Cancer Research* **65**, 2730-2737.

Martens U. M., Zijlmans M., Poon S., Dragowska W., Yui J., Chavez A., Ward K., Lansdorp M. (1998) *Nature Genetics* **18**, 76-80.

Mason J. M. and Biessmann H. (1995) The unusual telomeres of drosophila. *Trends in Genetics* **11**, 58-62.

Matsumoto Y. and Kim K. (1995) Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. *Science* **269**, 699-702.

Matsumoto Y., Kim K., Hurwitz J., Gary R., Levin D. S., Tomkinson A. E. and Park M. S. (1999) Reconstitution of proliferating cell nuclear antigen-dependent repair of apurinic/apyrimidinic sites with purified human proteins. *The Journal of Biological Chemistry* **274**, 33703-33708.

McClintock B. (1941) The stability of broken ends of chromosomes in *Zea Mays*. *Genetics* **26**, 234-282.

McEachern M. J. and Haber J. E. (2006) Break-induced replication and recombinational telomere elongation in yeast. *Annual Review of Biochemistry* **75**, 111-35.

Meyne J., Ratliff R. L. and Moyzis R. K. (1989) Conservation of the human telomere sequence (TTAGGG)n among vertebrates. *Proceedings of the National Academy of Sciences of the United States of Americal* **86**, 7049-7053.

Meister G., Landthaler M., Patkaniowska A., Dorsett Y., Teng G. and Tuschl T. (2004) Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Molecular Cell* **15**, 185-197.

Mertens F., Fletcher C. D., Dal Cin P., De Wever I., Mandahl N., Mitelman F., Rosai J., Rydholm A., Sciot R., Tallini G., Van den Berghe H., Vanni R. and Willen H. (1998) Cytogenetic analysis of 46 pleomorphic soft tissue sarcomas and correlation with morphologic and clinical features: a report of the CHAMP Study Group. Chromosomes and MorPhology. *Genes, Chromosomes & Cancer* **22**, 16-25.

Mitchell J. R., Cheng J. and Collins K. (1999) A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Molecular and Cellular Biology* **19**, 567-76.

Miyaki M., Konishi M., Tanaka K., Kikuchi-Yanoshita R., Muraoka M., Yasuno M., Igari T., Koike M., Chiba M. and Mori T. (1997) Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. *Nature Genetics* **17**, 271-272.

Modrich P. and Lahue R. (1996) Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annual Review of Biochemistry* **65**, 101-33.

Mondello C., Pirzio L., Azzalin C. M. and Giulotto E. (2000) Instability of interstitial telomeric sequences in the human genome. *Genomics* **68**, 111-117.

Montgomery E., Argani P., Hicks J. L., DeMarzo A. M. and Meeker A. K. (2004) Telomere lengths of translocation-associated and nontranslocation-associated sarcomas differ dramatically. *The American Journal of Pathology* **164**, 1523-1529.

Muller H. J. (1938) The remaking of chromosomes. *Collecting Net* **8**, 182-195.

Multani A. S. and Chang S. (2007) WRN at telomeres: implications for aging and cancer. *Journal of Cell Science* **120**, 713-721.

Muramatsu Y., Ohishi T., Sakamoto M., Tsuruo T. and Seimiya H. (2007) Cross-species difference in telomeric function of tankyrase 1. *Cancer Science* **98**, 850-857.

Murnane J. P., Sabatier L., Marder B. A. and Morgan W. F. (1994) Telomere dynamics in an immortal human cell line. *The EMBO Journal* **13**, 4953-62.

Nakagama H., Higuchi K., Tanaka E., Tsuchiya N., Nakashima K., Katahira M. and Fukuda H. (2006) Molecular mechanisms for maintenance of G-rich short tandem repeats capable of adopting G4 DNA structures. *Mutation Research* **598**, 120-131.

Natarajan S., Groff-Vindman C. and McEachern M. J. (2003) Factors influencing the recombinational expansion and spread of telomeric tandem arrays in *Kluyveromyces lactis*. *Eukaryotic Cell* **2**, 1115-1127.

Nicolaides N. C., Papadopoulos N., Liu B., Wei Y. F., Carter K. C., Ruben S. M., Rosen C. A., Haseltine W. A., Fleischmann R. D. and Fraser C. M. (1994) Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. *Nature* **371**, 75-80.

Nielsen F. C., Jager A. C., Lutzen A., Bundgaard J. R. and Rasmussen L. J. (2004) Characterization of human exonuclease 1 in complex with mismatch repair proteins, subcellular localization and association with PCNA. *Oncogene* **23**, 1457-1468.

Nieuwenhuis M. H. and Vasen H. F. (2007) Correlations between mutation site in APC and phenotype of familial adenomatous polyposis (FAP): a review of the literature. *Critical Reviews in Oncology/Hematology* **61**, 153-161.

Nowak M. A., Komarova N. L., Sengupta A., Jallepalli P. V., Shih I., Vogelstein B. and Lengauer C. (2002) The role of chromosomal instability in tumor initiation. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 16226-16231.

O'Connor M. S., Safari A., Xin H., Liu D. and Songyang Z. (2006) A critical role for TPP1 and TIN2 interaction in high-order telomeric complex assembly. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11874-11879.

Oda S., Maehara Y., Ikeda Y., Oki E., Egashira A., Okamura Y., Takahashi I., Kakeji Y., Sumiyoshi Y., Miyashita K., Yamada Y., Zhao Y., Hattori H., Taguchi K., Ikeuchi T., Tsuzuki T., Sekiguchi M., Karran P. and Yoshida M. A. (2005) Two modes of microsatellite instability in human cancer: differential connection of defective DNA mismatch repair to dinucleotide repeat instability. *Nucleic Acids Research* **33**, 1628-1636.

Okamura K., Ishizuka A., Siomi H. and Siomi M. C. (2004) Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes & Development* **18**, 1655-1666.

Olson A., Sheth N., Lee J. S., Hannon G. and Sachidanandam R. (2001) RNAi Codex: a portal/database for short-hairpin RNA (shRNA) gene-silencing constructs. *Nucleic Acids Research* **161; 34**, 25; D153-38; 7.

Opresko P. L., Mason P. A., Podell E. R., Lei M., Hickson I. D., Cech T. R. and Bohr V. A. (2005) POT1 stimulates RecQ helicases WRN and BLM to unwind telomeric DNA substrates. *The Journal of Biological Chemistry* **280**, 32069-32080.

Opresko P. L., von Kobbe C., Laine J. P., Harrigan J., Hickson I. D. and Bohr V. A. (2002) Telomere-binding protein TRF2 binds to and stimulates the Werner and Bloom syndrome helicases. *The Journal of Biological Chemistry* **277**, 41110-9.

Osterhage J. L., Talley J. M. and Friedman K. L. (2006) Proteasome-dependent degradation of Est1p regulates the cell cycle-restricted assembly of telomerase in *Saccharomyces cerevisiae*. *Nature Structural & Molecular Biology* **13**, 720-728.

Paddison P. J., Caudy A. A., Bernstein E., Hannon G. J. and Conklin D. S. (2002) Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes & Development* **16**, 948-958.

Paeschke K., Simonsson T., Postberg J., Rhodes D. and Lipps H. J. (2005) Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nature Structural & Molecular Biology* **12**, 847-854.

Palombo F., Iaccarino I., Nakajima E., Ikejima M., Shimada T. and Jiricny J. (1996) hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Current Biology* **6**, 1181-1184.

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

Pardue M. L., Rashkova S., Casacuberta E., DeBaryshe P. G., George J. A. and Traverse K. L. (2005) Two retrotransposons maintain telomeres in Drosophila. *Chromosome Research* **13**, 443-453.

Parsons R., Li G. M., Longley M., Modrich P., Liu B., Berk T., Hamilton S. R., Kinzler K. W. and Vogelstein B. (1995a) Mismatch repair deficiency in phenotypically normal human cells. *Science* **268**, 738-40.

Paule M. R. and White R. J. (2000) Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Research* **28**, 1283-1298.

Paull T. T. and Gellert M. (1998) The 3' to 5' exonuclease activity of Mre 11 facilitates repair of DNA double-strand breaks. *Molecular Cell* **1**, 969-979.

Pearson C. E., Nichol Edamura K. and Cleary J. D. (2005) Repeat instability: mechanisms of dynamic mutations. *Nature Reviews Genetics* **6**, 729-742.

Pedeutour F., Forus A., Coindre J. M., Berner J. M., Nicolo G., Michiels J. F., Terrier P., Ranchere-Vince D., Collin F., Myklebost O. and Turc-Carel C. (1999) Structure of the supernumerary ring and giant rod chromosomes in adipose tissue tumors. *Genes, Chromosomes & Cancer* **24**, 30-41.

Perrem K., Colgin L. M., Neumann A. A., Yeager T. R. and Reddel R. R. (2001) Coexistence of alternative lengthening of telomeres and telomerase in hTERT-transfected GM847 cells. *Molecular and Cellular Biology* **21**, 3862-75.

Petracek M. E., Lefebvre P. A., Silflow C. D. and Berman J. (1990) Chlamydomonas telomere sequences are A+T-rich but contain three consecutive G-C base pairs. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 8222-6.

Pickett H. A. (2002) Molecular characterisation of instability in human telomere repeat arrays. *PhD thesis*. University of Leicester.

Pickett H. A., Baird D. M., Hoff-Olsen P., Meling G. I., Rognum T. O., Shaw J., West K. P. and Royle N. J. (2004) Telomere instability detected in sporadic colon cancers, some showing mutations in a mismatch repair gene. *Oncogene* **23**, 3434-43.

Preston S. L., Wong W. M., Chan A. O., Poulsom R., Jeffery R., Goodlad R. A., Mandir N., Elia G., Novelli M., Bodmer W. F., Tomlinson I. P. and Wright N. A. (2003) Bottom-up histogenesis of colorectal adenomas: origin in the monocryptal adenoma and initial expansion by crypt fission. *Cancer Research* **63**, 3819-3825.

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.

Pryde F. E. and Louis E. J. (1999) Limitations of silencing at native yeast telomeres. *The EMBO Journal* **18**, 2538-50.

Qi H. and Zakian V. A. (2000) The Saccharomyces telomere-binding protein Cdc13p interacts with both the catalytic subunit of DNA polymerase alpha and the telomerase- associated est1 protein. *Genes & Development* **14**, 1777-88.

Radtke F. and Clevers H. (2005) Self-renewal and cancer of the gut: two sides of a coin. *Science* **307**, 1904-1909.

Rampino N., Yamamoto H., Ionov Y., Li Y., Sawai H., Reed J. C. and Perucho M. (1997) Somatic frameshift mutations in the BAX gene in colon cancers of the microsatellite mutator phenotype. *Science* **275**, 967-969.

Ramsden D. A. and Gellert M. (1998) Ku protein stimulates DNA end joining by mammalian DNA ligases: a direct role for Ku in repair of DNA double-strand breaks. *The EMBO Journal* **17**, 609-614.

Rashkova S., Karam S. E. and Pardue M. L. (2002) Element-specific localization of Drosophila retrotransposon Gag proteins occurs in both nucleus and cytoplasm. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 3621-3626.

Raynard S., Bussen W. and Sung P. (2006) A double Holliday junction dissolvasome comprising BLM, topoisomerase IIIalpha, and BLAP75. *The Journal of Biological Chemistry* **281**, 13861-13864.

Rayssiguier C., Thaler 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 Y. V., Ding Q., Lees-Miller S. P., Meek K. and Ramsden D. A. (2004) Non-homologous end joining requires that the DNA-PK complex undergo an autophosphorylation-dependent rearrangement at DNA ends. *The Journal of Biological Chemistry* **279**, 39408-39413.

Reenan R. A. and Kolodner R. D. (1992) Characterization of insertion mutations in the *Saccharomyces cerevisiae* MSH1 and MSH2 genes: evidence for separate mitochondrial and nuclear functions. *Genetics* **132**, 975-985.

Reichenbach P., Hoss M., Azzalin C. M., Nabholz M., Bucher P. and Lingner J. (2003) A human homolog of yeast Est1 associates with telomerase and uncaps chromosome ends when overexpressed. *Current Biology* **13**, 568-574.

Reya T. and Clevers H. (2005) Wnt signalling in stem cells and cancer. *Nature* **434**, 843-850.

Richards E. J. and Ausubel F. M. (1988) Isolation of a higher eukaryotic telomere from *Arabidopsis thaliana*. *Cell* **53**, 127-136.

Riethman H., Ambrosini A., Castaneda C., Finklestein J., Hu X. L., Mudunuri U., Paul S. and Wei J. (2004) Mapping and initial analysis of human subtelomeric sequence assemblies. *Genome Research* **14**, 18-28.

Riethman H. C., Xiang Z., Paul S., Morse E., Hu X.-L., Flint J., Chi H.-C., Grady D. L. and Moyzis R. K. (2001) Integration of telomere sequences with the draft human genome sequence. *Nature* **409**, 948-951.

Riou J. F., Guittat L., Mailliet P., Laoui A., Renou E., Petitgenet O., Megnin-Chanet F., Helene C. and Mergny J. L. (2002) Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 2672-2677.

Rivas F. V., Tolia N. H., Song J. J., Aragon J. P., Liu J., Hannon G. J. and Joshua-Tor L. (2005) Purified Argonaute2 and an siRNA form recombinant human RISC. *Nature Structural & Molecular Biology* **12**, 340-349.

Rizki A. and Lundblad V. (2001) Defects in mismatch repair promote telomerase-independent proliferation. *Nature* **411**, 713-6.

Ross-Macdonald P. and Roeder G. S. (1994) Mutation of a meiosis-specific MutS homolog decreases crossing over but not mismatch correction. *Cell* **79**, 1069-1080.

Royle N. J. (1995) The Proterminal Regions and Telomeres of Human Chromosomes. *Advances in Genetics* **32**, 273-315.

Sabah M., Cummins R., Leader M. and Kay E. (2006) Immunohistochemical detection of hTERT protein in soft tissue sarcomas: correlation with tumor grade. *Applied Immunohistochemistry & Molecular Morphology* **14**, 198-202.

Sambrook and Russell (2001) Molecular cloning, a laboratory manaul. Third edition, Cold Spring Harbor Laboratory Press, New York.

Samowitz W. S., Slattery M. L., Sweeney C., Herrick J., Wolff R. K. and Albertsen H. (2007) APC Mutations and Other Genetic and Epigenetic Changes in Colon Cancer. *Molecular Cancer Research* **5**, 165-170.

Sandberg A. A. (2004) Updates on the cytogenetics and molecular genetics of bone and soft tissue tumors: liposarcoma. *Cancer Genetics and Cytogenetics* **155**, 1-24.

Sasaki T., Shiohama A., Minoshima S. and Shimizu N. (2003) Identification of eight members of the Argonaute family in the human genome small star, filled. *Genomics* **82**, 323-330.

Sattler U., Frit P., Salles B. and Calsou P. (2003) Long-patch DNA repair synthesis during base excision repair in mammalian cells. *EMBO Reports* **4**, 363-367.

Scheel C., Schaefer K. L., Jauch A., Keller M., Wai D., Brinkschmidt C., van Valen F., Boecker W., Dockhorn-Dworniczak B. and Poremba C. (2001) Alternative lengthening of telomeres is associated with chromosomal instability in osteosarcomas. *Oncogene* **20**, 3835-44.

Schlotterer C. and Tautz D. (1992) Slippage synthesis of simple sequence DNA. *Nucleic Acids Research.* **20**, 211-215.

Schneider-Stock R., Boltze C., Jager V., Epplen J., Landt O., Peters B., Rys J. and Roessner A. (2003) Elevated telomerase activity, c-MYC-, and hTERT mRNA expression: association with tumour progression in malignant lipomatous tumours. *The Journal of Pathology* **199**, 517-525.

Schneider-Stock R., Walter H., Radig K., Rys J., Bosse A., Kuhnen C., Hoang-Vu C. and Roessner A. (1998) MDM2 amplification and loss of heterozygosity at Rb and p53 genes: no simultaneous alterations in the oncogenesis of liposarcomas. *Journal of Cancer Research and Clinical Oncology* **124**, 532-540.

Seidel C., Bartel F., Rastetter M., Bluemke K., Wurl P., Taubert H. and Dammann R. (2005) Alterations of cancer-related genes in soft tissue sarcomas: hypermethylation of RASSF1A is frequently detected in leiomyosarcoma and associated with poor prognosis in sarcoma. *International Journal of Cancer* **114**, 442-447.

Selva E. M., New L., Crouse G. F. and Lahue R. S. (1995) Mismatch correction acts as a barrier to homeologous recombination in *Saccharomyces cerevisiae*. *Genetics* **139**, 1175-1188.

Sfeir A. J., Chai W., Shay J. W. and Wright W. E. (2005) Telomere-end processing the terminal nucleotides of human chromosomes. *Molecular Cell* **18**, 131-138.

Shampay J., Szostak J. W. and Blackburn E. H. (1984) DNA sequences of telomeres maintained in yeast. *Nature* **310**, 154-157.

Sharma G. G., Gupta A., Wang H., Scherthan H., Dhar S., Gandhi V., Iliakis G., Shay J. W., Young C. S. and Pandita T. K. (2003) hTERT associates with human telomeres and enhances genomic stability and DNA repair. *Oncogene* **22**, 131-146.

Shell S. S., Putnam C. D. and Kolodner R. D. (2007) The N terminus of *Saccharomyces cerevisiae* Msh6 is an unstructured tether to PCNA. *Molecular Cell* **26**, 565-578.

Sheen F. M. and Levis R. W. (1994) Transposition of the line-like retrotransposon tart to drosophila chromosome termini. *Proceedings of the National Academy of Sciences of the United States of Americal* **91**, 12510-12514.

Shibata D., Peinado M. A., Ionov Y., Malkhosyan S. and Perucho M. (1994) Genomic instability in repeated sequences is an early somatic event in colorectal tumorigenesis that persists after transformation. *Nature Genetics* **6**, 273-281.

Shibata H., Toyama K., Shioya H., Ito M., Hirota M., Hasegawa S., Matsumoto H., Takano H., Akiyama T., Toyoshima K., Kanamaru R., Kanegae Y., Saito I., Nakamura Y., Shiba K. and Noda T. (1997) Rapid colorectal adenoma formation initiated by conditional targeting of the Apc gene. *Science* **278**, 120-123.

Shih I. M., Wang T. L., Traverso G., Romans K., Hamilton S. R., Ben-Sasson S., Kinzler K. W. and Vogelstein B. (2001) Top-down morphogenesis of colorectal tumors. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 2640-2645.

Singer M. S. and Gottschling D. E. (1994) TLC1: template RNA component of *Saccharomyces cerevisiae* telomerase. *Science* **266**, 404-409.

Smith L. L., Coller H. A. and Roberts J. M. (2003) Telomerase modulates expression of growth-controlling genes and enhances cell proliferation. *Nature Cell Biology* **5**, 474-479.

Smogorzewska A. and de Lange T. (2004) Regulation of telomerase by telomeric proteins. *Annual Review of Biochemistry* **73**, 177-208.

Smogorzewska A., Karlseder J., Holtgreve-Grez H., Jauch A. and de Lange T. (2002) DNA ligase IV-dependent NHEJ of deprotected mammalian telomeres in G1 and G2. *Current Biology* **12**, 1635-44.

Smogorzewska A., van Steensel B., Bianchi A., Oelmann S., Schaefer M. R., Schnapp G. and de Lange T. (2000) Control of human telomere length by TRF1 and TRF2. *Molecular and Cellular Biology* **20**, 1659-68.

Snow B. E., Erdmann N., Cruickshank J., Goldman H., Gill R. M., Robinson M. O. and Harrington L. (2003) Functional conservation of the telomerase protein Est1p in humans. *Current Biology* **13**, 698-704.

Snowden T., Acharya S., Butz C., Berardini M. and Fishel R. (2004) hMSH4-hMSH5 recognizes Holliday Junctions and forms a meiosis-specific sliding clamp that embraces homologous chromosomes. *Molecular Cell* **15**, 437-451.

Soda H., Raymond E., Sharma S., Lawrence R., Davidson K., Oka M., Kohno S., Izbicka E. and Von Hoff D. D. (2000) Effects of androgens on telomerase activity in normal and malignant prostate cells in vitro. *The Prostate* **43**, 161-168.

Song J. J., Smith S. K., Hannon G. J. and Joshua-Tor L. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434-1437.

Souza R. F., Appel R., Yin J., Wang S., Smolinski K. N., Abraham J. M., Zou T. T., Shi Y. Q., Lei J., Cottrell J., Cymes K., Biden K., Simms L., Leggett B., Lynch P. M., Frazier M., Powell S. M., Harpaz N., Sugimura H., Young J. and Meltzer S. J. (1996) Microsatellite instability in the insulin-like growth factor II receptor gene in gastrointestinal tumours. *Nature Genetics* **14**, 255-257.

Sparks A. B., Morin P. J., Vogelstein B. and Kinzler K. W. (1998) Mutational analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer. *Cancer Research* **58**, 1130-1134.

Stansel R. M., de Lange T. and Griffith J. D. (2001) T-loop assembly in vitro involves binding of TRF2 near the 3' telomeric overhang. *The EMBO Journal* **20**, 5532-40.

Starling J. A., Maule J., Hastie N. D. and Allshire R. C. (1990) Extensive telomere repeat arrays in mouse are hypervariable. *Nucleic Acids Research* **18**, 6881-6888.

Stasiak A. Z., Larquet E., Stasiak A., Muller S., Engel A., Van Dyck E., West S. C. and Egelman E. H. (2000) The human Rad52 protein exists as a heptameric ring. *Current Biology* **10**, 337-340.

Stavropoulos D. J., Bradshaw P. S., Li X., Pasic I., Truong K., Ikura M., Ungrin M. and Meyn M. S. (2002) The Bloom syndrome helicase BLM interacts with TRF2 in ALT cells and promotes telomeric DNA synthesis. *Human Molecular Genetics* **11**, 3135-44.

Stewart S. A., Hahn W. C., O'Connor B. F., Banner E. N., Lundberg A. S., Modha P., Mizuno H., Brooks M. W., Fleming M., Zimonjic D. B., Popescu N. C. and Weinberg R. A. (2002) Telomerase contributes to tumorigenesis by a telomere length-independent mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 12606-12611.

Strand M., Prolla T. A., Liskay R. M. and Petes T. D. (1993) Destabilization of tracts of simple repetitive dna in yeast by mutations affecting dna mismatch repair. *Nature* **365**, 274-276.

Sun X., Zheng L. and Shen B. (2002) Functional alterations of human exonuclease 1 mutants identified in atypical hereditary nonpolyposis colorectal cancer syndrome. *Cancer Research* **62**, 6026-6030.

Tabara H., Sarkissian M., Kelly W. G., Fleenor J., Grishok A., Timmons L., Fire A. and Mello C. C. (1999) The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* **99**, 123-132.

Takagi S., Kinouchi Y., Hiwatashi N., Nagashima F., Chida M., Takahashi S., Negoro K., Shimosegawa T. and Toyota T. (2000) Relationship between microsatellite instability and telomere shortening in colorectal cancer. *Diseases of the Colon and Rectum* **43**, S12-7.

Takai H., Smogorzewska A. and de Lange T. (2003) DNA damage foci at dysfunctional telomeres. *Current Biology* **13**, 1549-1556.

Takata H., Tanaka Y. and Matsuura A. (2005) Late S phase-specific recruitment of Mre11 complex triggers hierarchical assembly of telomere replication proteins in *Saccharomyces cerevisiae*. *Molecular Cell* **17**, 573-583.

Tarsounas M., Munoz P., Claas A., Smiraldo P. G., Pittman D. L., Blasco M. A. and West S. C. (2004) Telomere maintenance requires the RAD51D recombination/repair protein. *Cell* **117**, 337-47.

Teng S. C. and Zakian V. A. (1999) Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **19**, 8083-8093.

Tighe A., Johnson V. L. and Taylor S. S. (2004) Truncating APC mutations have dominant effects on proliferation, spindle checkpoint control, survival and chromosome stability. *Journal of Cell Science* **117**, 6339-6353.

Tokutake Y., Matsumoto T., Watanabe T., Maeda S., Tahara H., Sakamoto S., Niida H., Sugimoto M., Ide T. and Furuichi Y. (1998) Extra-chromosome telomere repeat DNA in telomerase-negative immortalized cell lines. *Biochemical and Biophysical Research Communications* **247**, 765-772.

Ulaner G. A., Hoffman A. R., Otero J., Huang H. Y., Zhao Z., Mazumdar M., Gorlick R., Meyers P., Healey J. H. and Ladanyi M. (2004) Divergent patterns of telomere maintenance mechanisms among human sarcomas: sharply contrasting prevalence of the alternative lengthening of telomeres mechanism in Ewing's sarcomas and osteosarcomas. *Genes, Chromosomes & Cancer* **41**, 155-162.

Ulaner G. A., Hu J. F., Vu T. H., Giudice L. C. and Hoffman A. R. (1998) Telomerase activity in human development is regulated by human telomerase reverse transcriptase (hTERT) transcription and by alternate splicing of hTERT transcripts. *Cancer Research* **58**, 4168-4172.

Ulaner G. A., Huang H. Y., Otero J., Zhao Z., Ben-Porat L., Satagopan J. M., Gorlick R., Meyers P., Healey J. H., Huvos A. G., Hoffman A. R. and Ladanyi M. (2003) Absence of a telomere maintenance mechanism as a favorable prognostic factor in patients with osteosarcoma. *Cancer Research* **63**, 1759-1763.

Umar A., Koi M., Risinger J. I., Glaab W. E., Tindall K. R., Kolodner R. D., Boland C. R., Barrett J. C. and Kunkel T. A. (1997) Correction of hypermutability, N-methyl-N'-nitro-N-nitrosoguanidine resistance, and defective DNA mismatch repair by introducing chromosome 2 into human tumor cells with mutations in MSH2 and MSH6. *Cancer Research* **57**, 3949-3955.

Umar A., Risinger J. I., Glaab W. E., Tindall K. R., Barrett J. C. and Kunkel T. A. (1998) Functional overlap in mismatch repair by human MSH3 and MSH6. *Genetics* **148**, 1637-1646.

Ushigome M., Ubagai T., Fukuda H., Tsuchiya N., Sugimura T., Takatsuka J. and Nakagama H. (2005) Up-regulation of hnRNP A1 gene in sporadic human colorectal cancers. *International Journal of Oncology* **26**, 635-640.

van Gent D. C., Hoeijmakers J. H. and Kanaar R. (2001) Chromosomal stability and the DNA double-stranded break connection. *Nature Review Genetics* **2**, 196-206.

van Steensel B. and de Lange T. (1997) Control of telomere length by the human telomeric protein TRF1. *Nature* **385**, 740-743.

van Steensel B., Smogorzewska A. and de Lange T. (1998) TRF2 protects human telomeres from end-to-end fusions. *Cell* **92**, 401-13.

Varley H., Pickett H. A., Foxon J. L., Reddel R. R. and Royle N. J. (2002) Molecular characterization of inter-telomere and intra-telomere mutations in human ALT cells. *Nature Genetics* **30**, 301-5.

Verdun R. E. and Karlseder J. (2006) The DNA damage machinery and homologous recombination pathway act consecutively to protect human telomeres. *Cell* **127**, 709-720.

Vogelstein B., Fearon E. R., Hamilton S. R., Kern S. E., Preisinger A. C., Leppert M., Nakamura Y., White R., Smits A. M. and Bos J. L. (1988) Genetic alterations during colorectal-tumor development. *The New England Journal of Medicine* **319**, 525-532.

Volker M., Mone M. J., Karmakar P., van Hoffen A., Schul W., Vermeulen W., Hoeijmakers J. H., van Driel R., van Zeeland A. A. and Mullenders L. H. (2001) Sequential assembly of the nucleotide excision repair factors in vivo. *Molecular Cell* **8**, 213-224.

Vulliamy T., Marrone A., Goldman F., Dearlove A., Bessler M., Mason P. J. and Dokal I. (2001) The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. *Nature* **413**, 432-5.

Waisberg M., Joseph P., Hale B. and Beyersmann D. (2003) Molecular and cellular mechanisms of cadmium carcinogenesis. *Toxicology* **192**, 95-117.

Walter M. F., Biessmann M. R., Benitez C., Torok T., Mason J. M. and Biessmann H. (2007) Effects of telomere length in Drosophila melanogaster on life span, fecundity, and fertility. *Chromosoma* **116**, 41-51.

Wang F., Podell E. R., Zaug A. J., Yang Y., Baciu P., Cech T. R. and Lei M. (2007) The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* **445**, 506-510.

Wang R. C., Smogorzewska A. and de Lange T. (2004) Homologous recombination generates T-loop-sized deletions at human telomeres. *Cell* **119**, 355-68.

Wang T. F., Kleckner N. and Hunter N. (1999) Functional specificity of MutL homologs in yeast: evidence for three Mlh1-based heterocomplexes with distinct roles during meiosis in recombination and mismatch correction. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 13914-13919.

Wang Y. and Qin J. (2003) MSH2 and ATR form a signaling module and regulate two branches of the damage response to DNA methylation. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15387-15392.

Watson P. and Lynch H. T. (1993) Extracolonic cancer in hereditary nonpolyposis colorectal cancer. *Cancer* **71**, 677-685.

Weber J. L. and Wong C. (1993) Mutation of human short tandem repeats. *Human Molecular Genetics* **2**, 1123-1128.

Wei K., Clark A. B., Wong E., Kane M. F., Mazur D. J., Parris T., Kolas N. K., Russell R., Hou H., Jr., Kneitz B., Yang G., Kunkel T. A., Kolodner R. D., Cohen P. E. and Edelmann W. (2003) Inactivation of Exonuclease 1 in mice results in DNA mismatch repair defects, increased cancer susceptibility, and male and female sterility. *Genes & Development* **17**, 603-614.

Wheeler J. M., Beck N. E., Kim H. C., Tomlinson I. P., Mortensen N. J. and Bodmer W. F. (1999) Mechanisms of inactivation of mismatch repair genes in human colorectal cancer cell lines: the predominant role of hMLH1. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 10296-301.

Wicky C., Villeneuve A. M., Lauper N., Codourey L., Tobler H. and Muller F. (1996) Telomeric repeats (TTAGGC)n are sufficient for chromosome capping function in Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 8983-8988.

Wightman B., Burglin T. R., Gatto J., Arasu P. and Ruvkun G. (1991) Negative regulatory sequences in the lin-14 3'-untranslated region are necessary to generate a temporal switch during Caenorhabditis elegans development. *Genes & Development* **5**, 1813-1824.

Woerner S. M., Kloor M., Mueller A., Rueschoff J., Friedrichs N., Buettner R., Buzello M., Kienle P., Knaebel H. P., Kunstmann E., Pagenstecher C., Schackert H. K., Moslein G., Vogelsang H., von Knebel Doeberitz M., Gebert J. F. and German H. C. (2005) Microsatellite instability of selective target genes in HNPCC-associated colon adenomas. *Oncogene* **24**, 2525-2535.

Wright W. E. and Shay J. W. (1995) Time, telomeres and tumors: is cellular senescence more than an anticancer mechanism. *Trends in Cell Biology* **5**, 293-297.

Wright W. E., Tesmer V. M., Huffman K. E., Levene S. D. and Shay J. W. (1997) Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes & Development* **11**, 2801-2809.

Wu G., Lee W. H. and Chen P. L. (2000) NBS1 and TRF1 colocalize at promyelocytic leukemia bodies during late S/G2 phases in immortalized telomerase-negative cells. Implication of NBS1 in alternative lengthening of telomeres. *The Journal of Biological Chemistry* **275**, 30618-22.

Wu L. and Hickson I. D. (2003) The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* **426**, 870-4.

Xin H., Liu D., Wan M., Safari A., Kim H., Sun W., O'Connor M. S. and Songyang Z. (2007) TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase. *Nature* **445**, 559-562.

Xu X., Peng M. and Fang Z. (2000) The direction of microsatellite mutations is dependent upon allele length. *Nature Genetics* **24**, 396-9.

Ye J. Z., Donigian J. R., van Overbeek M., Loayza D., Luo Y., Krutchinsky A. N., Chait B. T. and de Lange T. (2004) TIN2 binds TRF1 and TRF2 simultaneously and stabilizes the TRF2 complex on telomeres. *The Journal of Biological Chemistry* **279**, 47264-47271.

Yeager T. R. and Reddel R. R. (1999) Constructing immortalized human cell lines. *Current Opinion in Biotechnology* **10**, 465-9.

Zhang L., Yu J., Willson J. K., Markowitz S. D., Kinzler K. W. and Vogelstein B. (2001) Short mononucleotide repeat sequence variability in mismatch repair-deficient cancers. *Cancer Research* **61**, 3801-3805.

Zhang Q. S., Manche L., Xu R. M. and Krainer A. R. (2006) hnRNP A1 associates with telomere ends and stimulates telomerase activity. *RNA* **12**, 1116-1128.

Zhang Y., Yuan F., Presnell S. R., Tian K., Gao Y., Tomkinson A. E., Gu L. and Li G. M. (2005) Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell* **122**, 693-705.

Zhao J., Bilsland A., Jackson K. and Keith W. N. (2005) MDM2 negatively regulates the human telomerase RNA gene promoter. *BMC Cancer* **5**, 6.

Zhou X. P., Hoang J. M., Cottu P., Thomas G. and Hamelin R. (1997) Allelic profiles of mononucleotide repeat microsatellites in control individuals and in colorectal tumors with and without replication errors. *Oncogene* **15**, 1713-1718.

Zhu X. D., Kuster B., Mann M., Petrini J. H. and de Lange T. (2000) Cell-cycle-regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres. *Nature Genetics* **25**, 347-52.

Zhu X. D., Niedernhofer L., Kuster B., Mann M., Hoeijmakers J. H. and de Lange T. (2003) ERCC1/XPF removes the 3' overhang from uncapped telomeres and represses formation of telomeric DNA-containing double minute chromosomes. *Molecular Cell* **12**, 1489-98.