

HUMAN TELOMERES AND RECOMBINATION

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

by

Alberto Hidalgo Bravo MSc

Department of Genetics

University of Leicester

July 2012

HUMAN TELOMERES AND RECOMBINATION

Alberto Hidalgo Bravo

ABSTRACT

Telomeres are DNA-protein complexes that help protecting the end of linear chromosomes. They consist of repetitive DNA, in mammals the repeat unit is the hexanucleotide TTAGGG, these repeats span 5-20 kb. Under normal conditions in somatic cells, telomeres get shorter with every population doubling until they reach a critical length and then, the cell enters a checkpoint called senescence or M1 where it stops dividing. If the cell escapes senescence and continues dividing with further telomere shortening, it reaches a second checkpoint called crisis or M2. Crisis is characterized by telomere dysfunction leading to genomic instability that can end with cell death. However, some cells achieve to maintain telomere length by activating a telomere maintenance mechanism (TMM). The presence of a TMM is a hallmark of cancer cells. Two TMM have been described in human cells, one is the through the enzyme telomerase, which is active in 85% of cancers, and the second is a homologous recombination (HR) based mechanism called Alternative Lengthening of Telomeres (ALT) active in 15% of cancers. The evidence that the ALT pathway relies in HR was the observation that sequences can be copied from one telomere to another in ALT+ but not in telomerase+ cells and that several genes involved in HR are necessary for ALT progression.

The ALT pathway is not the only event involving HR at telomeres. It has been shown that the human herpesvirus 6 (HHV-6) can integrate into human telomeres. Interestingly, HHV-6 possesses perfect telomeric repeats within its genome. The proposed mechanism for integration is through HR between the telomeric repeats present in the virus with the human telomere repeats. The aim of this work is to unravel the molecular mechanism underlying the ALT pathway and HHV-6 integration. The data obtained will contribute to the understanding of HR in human telomeres.

ACKNOWLEDGEMENTS

I want to thank to Dr. Nicola J. Royle for the opportunity to work in her group and for her invaluable guidance and supervision. And thanks to the previous and current members of the laboratory G19 and G18 for their help.

I also want to offer my gratitude to the Mexican Council of Science and Technology (CONACYT) for funding my studies.

I want to dedicate this achievement to my beloved wife Lupita and also thank her for her unconditional love, support and company during these four years and the years to come.

Quiero agradecer a mi mama, mi papa por haberme enseñado y apoyado a conseguir mis metas, por todo su amor, comprensión y compañía en cada etapa de mi vida.

Thanks to my sisters and my brother because they have protected me and supported me during all these years.

At last but not least I want to thank to all my friends that have filled my life with good memories and joy.

CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	ii
LIST OF FIGURES	xi
LIST OF TABLES	xvi
ABREVIATIONS	xviii
CHAPTER 1 INTRODUCTION	1
1.1 Telomeres	1
1.2 Telomere structure	1
1.3 Telomere binding proteins	5
1.4 DNA damage and repair	6
1.4.1 DNA damage response	7
1.4.2 Double strand breaks repair	8
1.4.2.1 Non homologous end joining	9
1.4.2.2 Homologous recombination	11
1.4.2.3 Microhomology mediated end joining	13
1.5 Telomere dysfunction	15
1.6 Telomere maintenance mechanisms	16
1.6.1 Telomerase	16
1.6.1.1 Diseases associated with telomerase	18
1.6.2 The ALT mechanism	20
1.6.2.1 Hallmarks of ALT positive cells	20

1.6.2.2 Homologous Recombination and ALT models	23
1.7 RECQ Helicases	26
1.7.1 WRN and BLM helicases	27
1.7.2 WRN and BLM in telomere maintenance	28
1.8 The EXO1 exonuclease	30
1.8.1 EXO1 function in DSB repair by Homologous Recombination	33
1.8.2 EXO1 in Mismatch Repair	35
1.9 CtIP	39
1.9.1 CtIP and DSB repair	40
1.10 Human herpesvirus 6	42
1.10.1 Organization of HHV-6 genome	42
1.10.2 Clinical diseases associated to HHV-6 infection	44
1.10.3 HHV-6 Replication Cycle	46
1.10.4 Latent phase and chromosomal integration of HHV-6	48
1.10.5 Clinical implications of CIHHV-6	54
1.11 General aim	57
1.11.1 Specific aims	57
CHAPTER 2 MATERIAL AND METHODS	
Materials	58
2.1 Description of Single Telomere Length Analysis (STELA)	58
2.1.1 DNA preparation for STELA	60
2.1.2 Annealing Telorette2 to the genomic DNA prior to STELA	60
2.1.3 Size separation of PCR products using agarose gel electrophoresis	61
2.1.4 Re-amplification of STELA products	62
2.2 Southern Blot preparation and hybridization	62
2.2.1 Blotting and Hybridization	62
2.2.2 Post-hybridization washes	63

2.2.3 DNA probe labelling for Southern blot hybridization	64
2.3 Telomere Variant Repeat mapping by PCR (TVR-PCR)	64
2.3.1 Primer labelling	66
2.3.2 TVR-PCR	66
2.3.3 Polyacrylamide gel preparation	67
2.4 Human DNA extraction from cell lines	67
2.5 Polymerase chain reaction and sequencing	68
2.5.1 <i>EXO1</i>	68
2.5.2 HHV6	70
2.5.2.1 Chromosome-HHV6A junction PCR	70
2.5.2.2 PCR for HHV6 screening	70
2.6 Cell lines and cell culture	71
2.7 Protein extraction from cell pellet	73
2.8 Protein separation by PAGE and Western blot preparation	74
2.8.1 Protein detection using an antibody and staining	75
2.9 RNA Extraction	76
2.10 cDNA Synthesis	77
2.11 Real time PCR	77
2.11.1 Reference gene selection	78
2.11.2 qPCR efficiency	79
2.12 Inverse PCR (IPCR)	80
2.12.1 Southern blot preparation and hybridization to detect the chromosome-HHV6 junction fragment	82
2.12.2 Sequencing of the IPCR product	82
2.13 Telomere restriction fragment (TRF) length analysis	83
2.14 Transfection using pSUPERIOR vector	84
2.14.1 Vector	84
2.14.2 Design of short hairpin oligonucleotides	84
2.14.3 Vector cloning	86

2.14.4 Transformation of competent cells and plasmid purification.	87
2.14.5 Electroporation of human cell lines	89
2.14.6 Lipid-Base Transfection	89
2.14.7 Colony forming assay to evaluate the effect of CtIP and EXO1 downregulation	91
2.14.8 Single-cell colony isolation and expansion	92
2.14.9 Population doublings calculation	92
2.15 siRNA transfection	92
2.16 C-Circle assay	94
2.17 CtIP detection by Immunofluorescence and telomere detection using a Peptide Nucleic Acid probe	94
2.18 MS32 analysis	95
2.19 ss telomeric DNA assay	97

CHAPTER 3 ROLE OF THE WRN AND BLM HELICASES IN THE ALT PATHWAY

3.1 Background	99
3.2 Aim	102
3.3 Analysis of telomere mutations in the ALT+ cell line WV	102
3.4 TVR mapping on clones with downregulation of <i>BLM</i>	105
3.5 Discussion	107
3.6 Future work	110

CHAPTER 4 ROLE OF THE EXO1 EXONUCLEASE IN THE ALT PATHWAY

4.1 Background	111
4.2 Aim	113
4.3 Analysis of <i>EXO1</i> sequence in telomerase positive and ALT positive cell lines	113

4.4 Analysis of EXO1 expression by Western blot in telomerase+ and ALT cell lines	116
4.5 Expression analysis of EXO1 by real time PCR in telomerase+ and ALT cell lines	117
4.6 Downregulation of EXO1 in ALT positive cell lines	119
4.6.1 Colony forming assay in ALT cells tranfected with shRNA against <i>EXO1</i>	119
4.6.2 Expansion of clones after transfection with a shControl, shEXO1A or shEXO1B.	120
4.7 Discussion	126
4.8 Future work	128

CHAPTER 5 ROLE OF CtIP IN THE ALT PATHWAY

5.1 Background	130
5.2 Aim	132
5.3 Colocalization of CtIP with telomeres	132
5.4 Investigation of <i>CtIP</i> in ALT positive cell lines	135
5.4.1 Colony forming assay in ALT cells transfected with a vector expressing a shRNA against <i>CtIP</i>	135
5.4.2 Expansion of clones tranfected with a vector expressing shCtIPA, shCtIPB or shControl	136
5.5 Telomere analysis in a WI38V13/2RA clone shCtIPA (with downregulation of <i>CtIP</i>) using the STELA method	140
5.6 Effect of the downregulation of CtIP in the shCtIPA1 clone on markers of ALT activity	143
5.6.1 Effect of downregulation of <i>CtIP</i> on the MS32 minisatellite	143
5.6.2 Effect of downregulation of <i>CtIP</i> on the C-circle formation	145
5.6.3 Effect of downregulation of <i>CtIP</i> on telomere mutation	146

5.6.4 Detection of the telomeric 3' single strand overhang in WI38V13/2RA cell line	147
5.7 Effect of transient downregulation of <i>CtIP</i> on telomere length and C-circle formation on ALT cell lines	149
5.8 Discussion	154
5.9 Future work	158

**CHAPTER 6 CHROMOSOME INTEGRATED HUMAN HERPESVIRUS 6 (CIHHV-6)
MECHANISM AND CONSEQUENCES**

6.1 Background	161
6.2 Aim	163
6.3 CIHHV-6 Population screening	163
6.4 Telomere length analysis on chromosomes with CIHHV-6 from somatic and germline DNA	166
6.4.1 Telomere length analysis on chromosomes with CIHHV-6 from somatic DNA	166
6.4.2 Telomere length analysis on chromosomes with CIHHV-6 from germline DNA	171
6.5 Analysis of short telomeres on chromosomes with CIHHV-6 from somatic and germline DNA	174
6.6 Sequence analysis of the proximal and distal breakpoints on chromosomes with CIHHV-6	181
6.6.1 Sequence analysis of the proximal breakpoint on chromosomes with CIHHV-6	181
6.6.2 Sequence analysis of the distal breakpoint on chromosomes with CIHHV-6	188
6.7 Discussion	190
6.8 Future work	198

CHAPTER 7 FINAL DISCUSSION	201
7.1 The ALT pathway	202
7.2 Chromosomally integrated HHV-6	205
APPENDICES	
Appendix 1 Solutions and buffers	208
Appendix 2 pSUPERIOR plasmid	211
Appendix 3 Selection of reference genes for real time PCR	212
A3.1 Real time PCR efficiencies	214
Appendix 4 <i>EXO1</i> expression analysis in SUSM1 clones	217
Appendix 5 CtIP expression analysis in SUSM1 clones	218
Appendix 6 Analysis of the WI38V13/2RA clone shCtIPA1 and controls	219
A6.1 Southern blot of Single telomere length analysis (STELA) of the WI38V13/2RA clone shCtIPA1 and controls	219
A6.2 C-circle quantification in the WI38V13/2RA cell line, clone shCtIPA1 and controls	221
A6.3 Telomere mutation screening in the WI38V13/2RA shCtIPA1 clone by TVR mapping	221
A6.4 CtIP expression analysis by qPCR after transfection with siRNA against CtIP	222
A6.5 C-circle quantification in the ALT cell lines after transfection with siRNA against CtIP	223
Appendix 7 Analysis of samples with CIHHV-6	224

A7.1 Southern blots of STELAs in blood DNA from HHV-6 positive siblings used for telomere length analysis on telomeres XpYp, 12q, 17p and the HHV-6 associated telomere	224
A7.2 Southern blots of STELAs in sperm DNA from HHV-6 positive semen donors used for telomere length analysis on telomeres XpYp, 12q, 17p and the HHV-6 associated telomere	225
A7.3 Histograms of the data used for the telomere length analysis in the four sperm donors	229
A7.4 Southern blots of STELAs in sperm DNA from HHV-6 positive semen donors used for analysis of short telomeres on the XpYp, 12q, 17p and HHV-6 associated telomere	230
A7.5 Chromosome-HHV6-6 internal junction sequence from KUK	234
A7.6 Sequence of the distal chromosome-HHV-6 junction obtained from STELA products	235
A7.6.1 Sequence from three siblings with CIHHV-6A in 19q	235
A7.6.2 Sequence from sperm donor 1 harbouring CIHHV-6A	236
A7.6.3 Sequence from sperm donor 56 harbouring CIHHV-6B	237
BIBLIOGRAPHY	238

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.1 Structure of human telomeres	2
1.2 The “end replication problem”	4
1.3. Telomere length maintenance and immortalization	4
1.4. Representation of the shelterin complex	6
1.5. DSB repair by Non homologous end joining (NHEJ)	10
1.6. Models for DSB repair by HR	13
1.7 Microhomology mediated end joining (MMEJ)	15
1.8 Proposed models for telomere elongation in ALT cells	25
1.9. EXO1 functional domains	33
1.10 Functional domains of EXO1 and localization of the missense mutations identified	36
1.11. HHV-6 genome organization	43
1.12 Simplified diagram of the lytic cycle of human herpesviruses	47
1.13. Orientation of the Chromosomal integrated HHV-6 (CIHHV-6)	49
1.14. Proposed model for CIHHV-6	53

CHAPTER 2 MATERIAL AND METHODS

2.1. The STELA technique	60
2.2. TVR-PCR technique	65
2.3. IPCR	81
2.4 Diagram showing the annealed pair of oligonucleotides used to knock down the CtIP mRNA	85

CHAPTER 3 ROLE OF THE WRN AND BLM HELICASES IN THE ALT PATHWAY

3.1 XpYp STELA products and TVR maps from the W-V cell line and clone 6	104
---	-----

3.2 Telomere Variant Repeats maps from the progenitor allele of the XpYp telomere in the WV cell line	106
---	-----

CHAPTER 4 ROLE OF THE EXO1 EXONUCLEASE IN THE ALT PATHWAY

4.1. Known SNPs in the coding sequence of EXO1	114
4.2 EXO1 Western blot	117
4.3 Relative expression of EXO1	118
4.4 Colony forming assay in ALT+ cell lines transfected with shEXO1 plasmids	120
4.5 Growth curves of the clones expanded after transfection of the ALT+ cell lines	121
4.6 EXO1 qPCR expression analysis	123
4.7 Comparison of the EXO1 expression analysis	125

CHAPTER 5 ROLE OF CtIP IN THE ALT PATHWAY

5.1 Model proposed for C-strand processing at telomeres in yeast after DNA replication	131
5.2 CtIP immunofluorescence in cell lines	134
5.3 Colony forming assay following transfection with shCtIPA or shCtIPB plasmid	136
5.4 Growth curves of the expanded clones from the ALT cell lines	137
5.5 Expression analysis of CtIP by qPCR	140
5.6 Length analysis by STELA of the XpYp telomere of clone shCtIPA1 and control clones	142
5.7 Analysis of STELA products from telomeres 12q and 17p obtained from the shCtIPA1 and four control clones	142
5.8 TRF of WI38V13/2RA cell line and clones	143
5.9 Southern blots of the gels used to calculate MS32 mutation rate on the WIV13/2RA shCtIPA1 clone	144
5.10 Quantification of telomeric DNA generated in the C-circle assay from the WI38V13/2RA cell line, shCtIPA1 and control clones.	145
5.11 XpYp telomere maps for the WI38V13/2RA cell line	147

5.13 Detection of single stranded (TTAGGG) _n DNA by dot blot	149
5.14 qPCR analysis of CtIP expression after transfection with siCtIP or siControl	150
5.15 TRF of ALT+ and the telomerase+ cell lines.	151
5.16 Quantification of the C-circle detection assay on cell lines after transfection with siCtIP or siControl	152
5.17 qPCR analysis of CtIP expression after transfection with siCtIP or siControl.	153
5.18 Quantification of C-circle DNA in ALT+ cell lines after transfection with siCtIP or siControl	154

CHAPTER 6 CHROMOSOME INTEGRATED HUMAN HERPESVIRUS 6 (CIHHV-6)

MECHANISM AND CONSEQUENCES

6.1 STELA on HHV-6	164
6.2 Examples of STELA Southern blots from siblings 1499, 1500 and 1501 used for telomere length analysis	168
6.3 Telomere length analysis from blood DNA from three siblings carrying CIHHV6-A at 19q	169
6.4 STELA Southern blot analysis in sperm DNA from donor 1	172
6.5 Telomere length analysis on DNA from germline	173
6.6 Analysis of telomeres below 2.33SD from the mean	177
6.7 STELA Southern blot analysis in sperm DNA from donor 1	179
6.8 Analysis of short telomeres (≤ 1.5 kb) found in DNA from the four sperm donors	181
6.9. Internal junction of HHV-6A and human chromosome 10q was obtained by IPCR from KUK cell line	183
6.10 STELA Southern blot analysis in DNA from the KUK, HT1080 and AMD cell lines using the primer 10q	185
6.11 Southern blots of the IPCR products from the AS, AMD, OL and MP cell lines harbouring CIHHV-6B	186

6.12 Alignment of the sequence obtained from DR8 region from the HHV-6B strain integrated in the lymphoblastoid cell line MP with the consensus sequence of the strain HST of HHV-6B	187
6.13 Agarose gel showing a PCR on genomic DNA from 18 different lymphoblastoid cell lines with CIHHV-6	188
6.14 Schematic representations of the sequence organisation obtained from STELA products of the CIHHV-6 samples	190
6.15 Possible mechanisms for CIHHV-6 reactivation	198

APPENDICES

A.1 Plasmid pSUPERIOR.neo+GFP	211
A.2 Determination of the reference genes	212
A.3 Standard curves to calculate the amplification efficiency of the β -Actin, GAPDH, EXO1, CtIP, UBC, SDHA and ATP5B amplicons and comparison of the amplification efficiency between amplicons	214
A.4 Replicas of the EXO1 expression analysis by qPCR in the SUSM1 clones	216
A.5 Replica of the CtIP expression analysis by qPCR in the SUSM1 clones	218
A.6 Southern blot of XpYp STELA from WI38V13/2RA control clones	219
A.7 Southern blot of STELA from WI38V13/2RA shCtIPA1 and control clones	220
A.8 Replica of the quantification of telomeric DNA generated in the C-circle assay	221
A.9 Telomere mutation screening by TVR mapping in the WI38V13/2RA clone shCtIPA1	221
A.10 Technical replica of qPCR analysis of CtIP expression after transfection with siCtIP or siControl	222
A.11 Technical replicates of the qPCR analysis of CtIP expression after transfection with siCtIP or siControl.	222
A.12 Replica of the quantification of C-circle DNA ALT+ cell lines after the second transfection with siCtIP or siControl	223
A.13 STELA Southern blot analysis in blood DNA from the HHV-6 positive siblings	224

A.14 STELA Southern blot analysis in sperm DNA from the HHV-6 positive donors used for telomere length analysis.	225
A.15 Histograms generated from the data used in the telomere length analysis from the four sperm donors	226
A.16 STELA Southern blot analysis in sperm DNA from the HHV-6 positive donors used for analysis of short telomeres.	227
A.17 Complete sequence of the Chromosome-HHV-6AA internal junction from the lymphoblastoid cell line obtained by IPCR.	234
A.18 Sequence of the distal chromosome-HHV-6A junction obtained from STELA products from blood DNA from three siblings	235
A.19 Sequence of the distal chromosome-HHV-6A junction obtained from STELA products from sperm DNA from donor 1 harbouring CIHHV-6A.	236
A.20 Sequence of the distal chromosome-HHV-6A junction obtained from STELA products from sperm DNA from donor 56 harbouring CIHHV-6B	237

LIST OF TABLES

CHAPTER 1 INTRODUCTION

1.1 Diseases associated with telomere biology dysfunctions and genes involved	19
1.2 Position, length and function of the domains of the human EXO1 protein	32
1.3 Variants found in the sequence of <i>EXO1</i>	38

CHAPTER 2 MATERIAL AND METHODS

2.1 Primers used for single telomere amplification (STELA)	59
2.2 TVR primers	65
2.3 Primers used for telomere variant repeat (TVR) mapping	66
2.4 Primers used for EXO1 PCR amplification and sequencing	69
2.5 Primers used for HHV-6 PCRs	71
2.6 Cell lines used and medium used for each cell line	72
2.7. Antibodies used for Western blot analysis	75
2.8 Genes included in the GeNorm reference genes selection kit from Primer design	79
2.9 Primers used for EXO1 and CtIP real time PCR	79
2.10 Primers used for inverse (PCR IPCR)	80
2.11 Sequence of the oligonucleotides used to generate the shRNA against CtIP and EXO1	86
2.12 Lipid based transfection	90
2.13 Primers used for amplification of the MS32 minisatellite	96

CHAPTER 3 ROLE OF THE WRN AND BLM HELICASES IN THE ALT PATHWAY

3.1 Mutation frequencies and rates for the groups studied	107
---	-----

CHAPTER 4 ROLE OF THE EXO1 EXONUCLEASE IN THE ALT PATHWAY

4.1 SNPs found in EXO1 in eight different cell lines	115
--	-----

CHAPTER 5 ROLE OF CtIP IN THE ALT PATHWAY

5.1 Cell lines were screened for colocalization of CtIP and telomeric DNA	133
---	-----

CHAPTER 6 CHROMOSOME INTEGRATED HUMAN HERPESVIRUS 6 (CIHHV-6)

MECHANISM AND CONSEQUENCES

6.1 CIHHV-6 frequencies in the HDGP-CEPH panel and Spanish samples	165
6.2 Frequency of CIHHV-6 in a panel of 92 DNA samples from sperm from British donors	166
6.3 Comparison of mean telomere length between the CIHHV-6A telomere (DR1R) and 12q, 17p and XpYp telomeres from the three sibling	170
6.4 Analysis of telomeres below 2.33SD from the mean telomere length	176
6.5 Fisher's exact test results	177
6.6 Analysis of short telomeres	180
6.7 Fisher's exact test results	180
6.8 Known location of the CIHHV-6 in the lymphoblastoid telomerase+ EBV transformed cell lines from five unrelated donors	182

ABBREVIATIONS

Bp	Base pair
cDNA	Complementary DNA
DNA	Desoxyribonucleic acid
Kb	Kilo base
KDa	Kilo Daltons
Mb	Mega base
MEM	Minimal essential medium
mRNA	Messenger RNA
μ F	Micro Faraday
μ g	Micrograms
mg	Milligrams
ml	Millilitre
μ l	Microlitre
μ M	Micromolar
mM	Milimolar
M	Molar
Ng	Nanograms
PCR	Polymerase chain reaction
Pg	Picograms
Pmol	Picomols
RNA	Ribonucleic acid
RPM	Revolutions per minute

CHAPTER 1 - INTRODUCTION

1.1 Telomeres

In the decade of the 1940's Hermann Müller was the first to describe what he called telomeres, in the fruit fly, *Drosophila melanogaster*. Müller noted that chromosomal inversions derived from ionizing radiation-induced double strand breaks (DSB) never involved the original ends of the chromosomes. The word telomere comes from *telos*, which in Greek means "end", and *meros*, means "part". A few years later Barbara McClintock noticed that while broken ends of maize chromosomes were capable of fusion, this was rarely observed with unbroken chromosomes, this suggested a distinct functional nature of intact ends (De Boeck G *et al.* 2009).

In 1961, Hayflick and Moorhead observed that normal somatic cells have a limited capacity to divide, after which they become senescent; this was called the Hayflick limit (Hayflick & Moorhead 1961). Ten years later Olovnikov hypothesized that this could result from the incomplete replication at the ends of linear DNA molecules (Olovnikov AM 1973). In the same decade Blackburn and Gall cloned the first telomere DNA structure from the ciliated protozoan *Tetrahymena thermophila*, and found a repetitive G-rich sequence, TTGGGG (Blackburn EH *et al.* 1978). By the end of the 1980's the human telomere sequence was identified by Moyzis RK *et al.* (Moyzis RK *et al.* 1988).

1.2 Telomere structure

Telomeres are DNA-protein complexes at the ends of linear eukaryotic chromosomes; they help ensure genomic stability by protecting the ends of the chromosomes from degradation or random fusion events, and by facilitating telomere elongation to compensate for replicative erosion (Palm W *et al.* 2008). Telomeres consist of long tracts of repetitive DNA sequence that in humans range from 5 to 20 kb in germline in length,

but in mice they can range from 20 to 150 kb (Palm W *et al.* 2008). One characteristic of the telomere is that the strand of the 3' end is rich in guanosine and is called the G strand and consequently the complementary strand is called the C strand. All vertebrates use the sequence (TTAGGG)_n at their chromosome ends (Griffith JD *et al.* 1999); the telomeric region of human chromosomes has some degenerate repeats (e.g. TCAGGG) at the beginning of the telomere. The telomeres are not blunt-ended; they have a protrusion of the G strand of 50 to 500 nucleotides known as the 3' overhang. This overhang is produced by the end replication problem and through post replication resection events (Palm W *et al.* 2008). *In vivo* studies have demonstrated that the double-stranded telomeric region can bend around allowing a single stranded terminus to invade into the double stranded region, thus forming a D-loop. The whole structure is referred to as a t-loop (Telomere-loop) (Griffith JD *et al.* 1999) (Figure 1.1). Both elements, the single stranded DNA and the repetitive nature of telomeres, are critical for maintaining t-loops. The t-loop is useful to hide the chromosome ends from the DNA repair machinery (Compton SA *et al.* 2007).

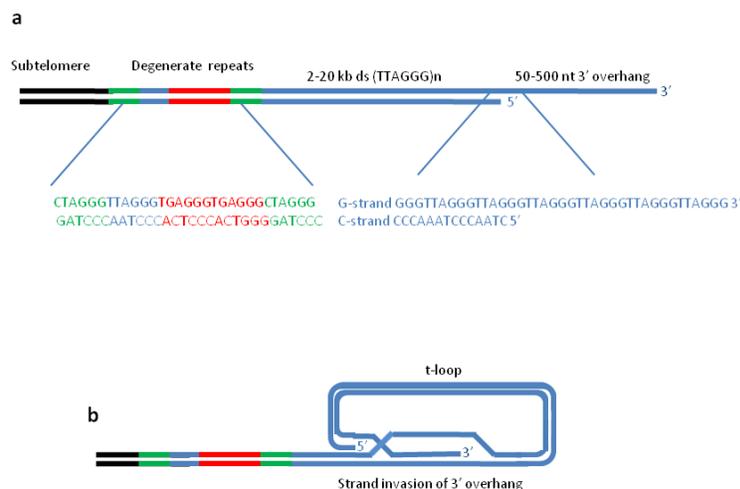


Figure 1.1 Structure of human telomeres. (a) Human chromosome end in an array of TTAGGG repeats. At the beginning of the telomere there is a region of degenerate telomere repeats. The telomere terminus contains a long 3' single stranded overhang on the G strand. The 5' end of human chromosomes nearly always finishes with the sequence ATC-5'. (b) Schematic representation of the t-loop, the 3' single stranded overhang folds to invade the duplex DNA and anneal with the C-strand so creating a displacement loop (d-loop) of variable size. Modified from (Palm W *et al.* 2008).

In normal somatic cells, telomeres shorten by 50 to 200 base pairs per cell division because of the end replication problem and the processing of the 5' end at the leading strand in order to generate the 3' overhang (Harley CB *et al.* 1990) (Figure 1.2). It is thought that this telomere shortening process acts like a "mitotic clock" limiting the number of times a cell can divide, and works as an important tumor suppressor mechanism (Reddel RR 2003). When the telomere length declines below a certain threshold, an irreversible growth arrest state, referred to as replicative senescence or "mortality stage 1" (M1), is triggered. In this state cells cease to divide but remain alive and metabolically active (Wei W *et al.* 1999). It is known that human cells can bypass M1 senescence by the inactivation of the p53 and Rb tumor suppressor pathways allowing cells to escape from senescence, this results in continued cellular division and further telomere shortening (Reddel RR 2003). These cells eventually reach a second proliferative block referred to as crisis (M2), which is characterized by telomere dysfunction and cell death. Cells emerging from crisis have to activate a telomere maintenance mechanism (TMM) to avoid cell death. In human cells there are at least two possible TMMs, the activation of the enzyme telomerase or the poorly understood mechanism of telomere maintenance referred to as ALT (Alternative Lengthening of Telomeres) (Shay JW *et al.* 2004) (Figure 1.3), this leads to cellular immortalization: a hallmark of human cancer. Approximately 85% of human cancers are telomerase-positive (Shay JW *et al.* 1997), many of the remaining 15% of cancers maintain their telomeres by the ALT mechanism (Bryan TM *et al.* 1997).

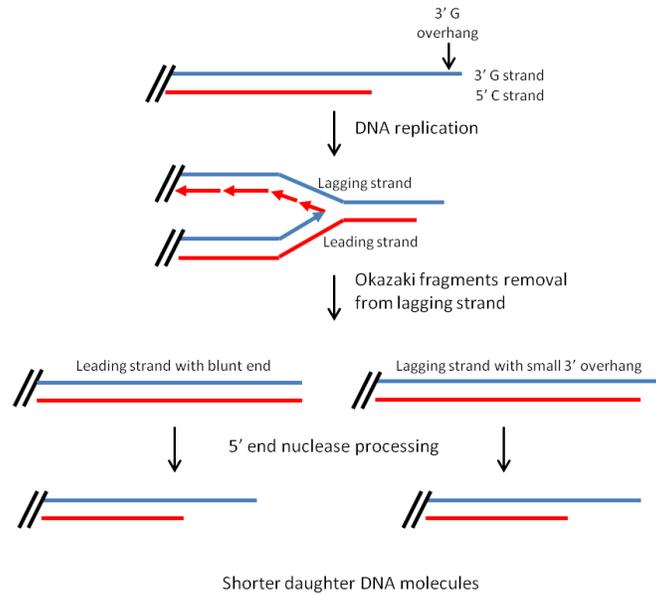


Figure 1.2 The “end replication problem” represents the inability of semiconservative DNA replication to copy the full length of linear DNA molecules. Removal of the RNA primer from the last Okazaki fragment in the lagging strand generates a short 3’ single stranded overhang. Further processing of the 5’ end by exonucleases is needed to generate a longer 3’ overhang for t-loop formation. On the other hand the leading strand finishes in a blunt end, which has to be processed by 5’ exonucleases to generate the 3’ single stranded overhang. This process results in daughter molecules having shorter telomeres than the progenitor molecule.

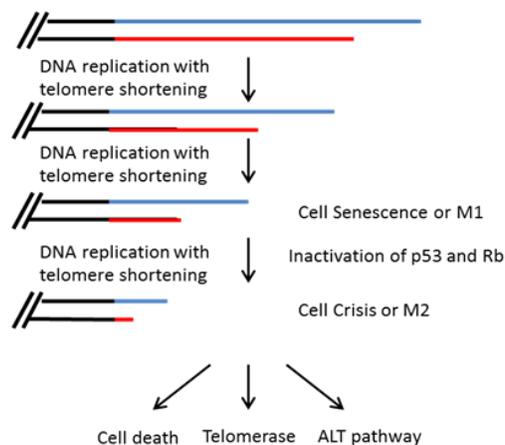


Figure 1.3. Telomere length maintenance and immortalization. With each population doubling the telomere becomes shorter leading the cell to senescence. If the cell escapes senescence and continues dividing, it will reach a critical length of telomeres and the cell will enter into the crisis stage. In order to bypass crisis, it is necessary to activate a Telomere Maintenance Mechanism (TMM), either telomerase expression or ALT, otherwise the cell will die.

1.3 Telomere binding proteins

The telomeric DNA in mammals is associated with six telomere specific proteins which form the shelterin complex and consist of: the telomere-repeat-binding factors 1 and 2 (TRF1 and TRF2), the TRF1 interacting protein 2 (TIN2), the protection of telomere protein 1 (POT1), the transcriptional repressor/activator protein 1 (RAP1), and the TIN2 and POT1 organizing protein (TPP1) (de Lange T 2005). The 3' overhang of telomeres could be recognised as an intermediate of double strand break (DSB) repair, and could lead to chromosome fusions and/or degradation and cell cycle arrest. Shelterin enables the cell to distinguish the natural ends of chromosomes from DNA (double strand) breaks and repress DNA damage response (Palm W *et al.* 2008).

TRF1 and TRF2 bind to telomeric double-stranded (dsDNA) as homodimers or oligomers (as shown in Figure 1.4), however they do not interact with each other directly, RAP1 forms a 1:1 complex with TRF2, TIN2 can interact with TRF1, TRF2 and TPP1 (Palm W *et al.* 2008). This acts as a bridge between the proteins bound to the dsDNA and to the single-stranded DNA (ssDNA). TPP1 is recruited by TIN2 and also interacts with POT1 whilst POT1 binds specifically to the G-rich single stranded telomeric DNA. In addition to the shelterin complex, other proteins bind to the telomeric chromatin, but unlike shelterin, they are not present at the telomere throughout the whole cell cycle and they have roles at other sites in the genome. These proteins are involved in different cellular process such as DNA damage signalling, DNA repair or chromatin organisation. As a result these proteins are referred to as shelterin accessory factors (Palm W *et al.* 2008) (Figure 1.4).

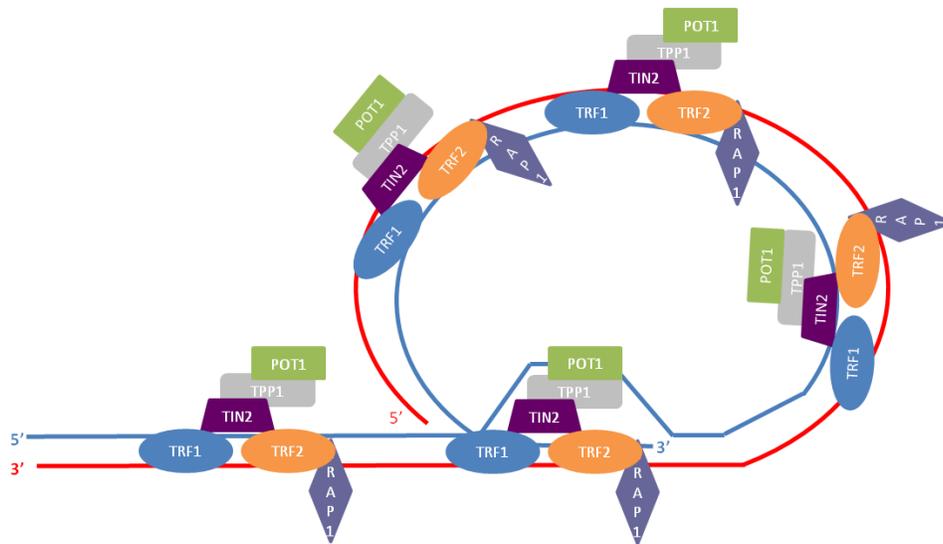


Figure 1.4. Representation of the shelterin complex. TRF1 and TRF2 bind to double-stranded (dsDNA) as homodimers or oligomers. RAP1 forms a 1:1 complex with TRF2. TIN2 can interact with TRF1, TRF2 and TPP1, it acts as a bridge between the proteins bound to the dsDNA and to the single-stranded DNA (ssDNA). TPP1 is recruited by TIN2 and also interacts with POT1 which specifically binds to the G-rich single stranded telomeric DNA. Modified from (De Boeck G *et al.* 2009).

Telomere length, the shelterin complex and the proteins involved in DNA recombination/repair, such as the BLM and WRN helicases, are very important factors to insure telomere function (Loayza D *et al.* 2003). It has been proposed that telomere attrition results in insufficient shelterin binding and failure to repress the activation of the DNA damage signal via the Ataxia Telangiectasia Mutated (ATM) and Ataxia Telangiectasia and RAD3 Related (ATR) protein kinases. This may result in the progression of the damage response and cell arrest (Loayza D *et al.* 2003).

1.4 DNA damage and repair

DNA is constantly damaged by exogenous factors such as ionizing radiation or chemicals, but also by endogenous factors sometimes related to cellular process such as oxidative damage from reactive oxygen species produced in the mitochondria. If the DNA damage is

not adequately repaired it can cause stalling of replication forks and deleterious mutations during DNA replication (Germann MW *et al.* 2010). The effects of unrepaired DNA damage can be devastating. When DNA damage accumulates it can lead to mutations in key genes of essential cellular pathways, activation of oncogenes or inactivation of tumor suppressor genes. These events can end in cell death or cell transformation. In order to maintain genomic stability the cell has developed a variety of mechanism for repairing the DNA damage. The group of these mechanisms is referred as DNA damage response (DDR). DDR has evolved in different, but interconnected pathways, for specific types of DNA damage. The general scheme for the majority of DDR pathways includes detection of the DNA damage, recruitment of DNA repair factors at the site of the damage and the physical repair of the lesion (Lord CJ *et al.* 2012). Below is a brief description of DDR mechanisms, with the main focus on double strand breaks (DSBs) repair.

1.4.1 DNA damage response

DNA can be altered by modifications caused by oxidative reagents, alkylating agents or single strand breaks (SSBs), and these kinds of lesion are repaired by the base excision repair (BER) pathway (Lord CJ *et al.* 2012). The BER pathway comprises removal of the damaged bases from the DNA, followed by excision of DNA backbone at the site of the lesion, and finally DNA synthesis. Two key participants in BER, acting as damage sensors and signal transducers, are the poly (ADP-ribose) polymerase 1 and 2 (PARP1 and PARP2) enzymes (Lord CJ *et al.* 2012). Some other lesions, like thymidine dimers caused by ultraviolet light, generate a major distortion of the DNA structure and these lesions are repaired by the nucleotide excision repair (NER) pathway (Lord CJ *et al.* 2012). NER is frequently subdivided into transcription-coupled NER and global-genome NER. Transcription-coupled NER is carried out when the lesion is blocking transcription progression and is detected by the RNA polymerase (Lord CJ *et al.* 2012). Global-genome NER is when the lesion is not detected by the transcription machinery but it disturbs base

pairing and DNA structure. Excision repair cross-complementing protein 1 (ERCC1) is the key participant of this pathway (Lord CJ *et al.* 2012). ERCC1 forms a heterodimer with xeroderma pigmentosum group F (XPF), this heterodimer has endonuclease activity which is essential for incising the DNA at the site of damage (Rageul J *et al.* 2011). Sometimes during DNA replication nucleotides are incorporated erroneously by the DNA polymerase causing mismatches. The mismatched bases disrupt the DNA double helix structure and are recognized by the mismatch repair (MMR) system. After recognition of the damage, the MMR system excises the DNA on both sides of the lesion followed by new DNA synthesis. In mammals, the MMR system is formed by the proteins MSH2, MSH3, MSH6, MLH1, MLH3, PMS1 and PMS2 (Lord CJ *et al.* 2012).

1.4.2 Double strand break repair

One of the most serious lesions in DNA are DSBs. DSB repair is critical for genome integrity and cell survival. DSBs can be caused by external agents like ionizing radiation or as a result of normal cellular processes like meiosis. There are two main mechanisms for DSB repair: non-homologous end joining (NHEJ) and homologous recombination (HR) (Symington LS *et al.* 2011). Recently a third mechanism has been identified, called Microhomology mediated end joining (MMEJ). The selection of the repair mechanism depends on the structure of the DNA ends and the phase of the cell cycle (Symington LS *et al.* 2011).

After DSBs occur, the MRN/X complex (Mre11-Rad50-MBS1/Xrs2) is rapidly recruited to the sites of damage to initiate the signalling cascade and cell cycle arrest (Shiloh Y 2003). The MRN complex recruits the phosphatidylinositol 3-kinase-related protein kinase ATM and ATR to the site. It seems that DSBs trigger the activation of ATM, whereas ATR activation requires the formation of ssDNA (Shiloh Y 2003). As part of the signalling, the histone H2AX is phosphorylated (γ H2AX) by both kinases on Serine 139, over a large chromatin domain surrounding the site of a DSB and promotes the local accumulation of

other DNA damage response factors, resulting in cytologically detectable foci that span hundreds of kb (Shiloh Y 2003). Even though the mechanism is not fully understood, these DNA damage foci have been implicated in signal amplification and DNA repair. ATM and ATR can further phosphorylate two kinases, Chk1 and Chk2, both able to block cell cycle progression in G1/S or G2/M, through the inhibition of the Cdc25 phosphatases, which are required for the activation of Cdks. Chk1 and Chk2 also cooperate with ATM and ATR to activate p53, which further inhibits cell cycle progression by the induction of the Cdk inhibitor p21 (Palm W *et al.* 2008). Not only after the DDR has been initiated, the cell can engage with one of the possible DSB repair mechanisms. As mentioned above the DNA structure and the cell cycle phase influence in the pathway chosen, however the entire process underlying this decision is not fully understood.

1.4.2.1 Non homologous end joining

NHEJ is the simplest mechanism to repair DSBs and can operate during the entire cell cycle. It is more prevalent on G1 phase (Takata M *et al.* 1998). The main participants are the proteins Ku70 and Ku 80, which form a heterodimer (Ku). Ku binds to the broken ends of the DSB in conjunction with the MRN complex, and brings together the blunt ends of the DSB to protect them from degradation. It is very important to protect the DSB ends from excessive processing since the NHEJ mechanism cannot be carried out when the long single strand overhangs are generated. By avoiding the generation of long single stranded overhangs at the DSB, Ku directs the DSB repair towards NHEJ (Feldmann E *et al.* 2000). However, when the nucleotides at the ends of the DSB are damaged they need to be processed in order to ligate them. The MRN complex and Ku are also necessary to recruit, stabilise and stimulate the activity of the DNA ligase IV and its accessory proteins, XRCC4 and XLF (XRCC4-like factor) (Ahnesorg P *et al.* 2006). Before ligation can be carried out, the ends of the DSB must have ligatable 5' phosphates and 3' hydroxyl group. If the 5' or 3' ends are damaged and not available for ligation, the XRCC4 factor is able to recruit

polynucleotide kinase (PNK). PNK contains 5' kinase and 3' phosphatase activities, which make it capable to perform the modifications needed to generate “ligatable ends” (Chappell C *et al.* 2002). In addition, if damaged or mismatched bases are present, the activity of one nuclease may be needed. Also the ends of the DSB may need to be processed by a nuclease to allow their pairing. To finish the repair of the DSB DNA polymerases are recruited to fill the single-stranded gaps (Figure 1.5). The DNA polymerases used for NHEJ are the Pol λ and Pol μ (Pardo B *et al.* 2009).

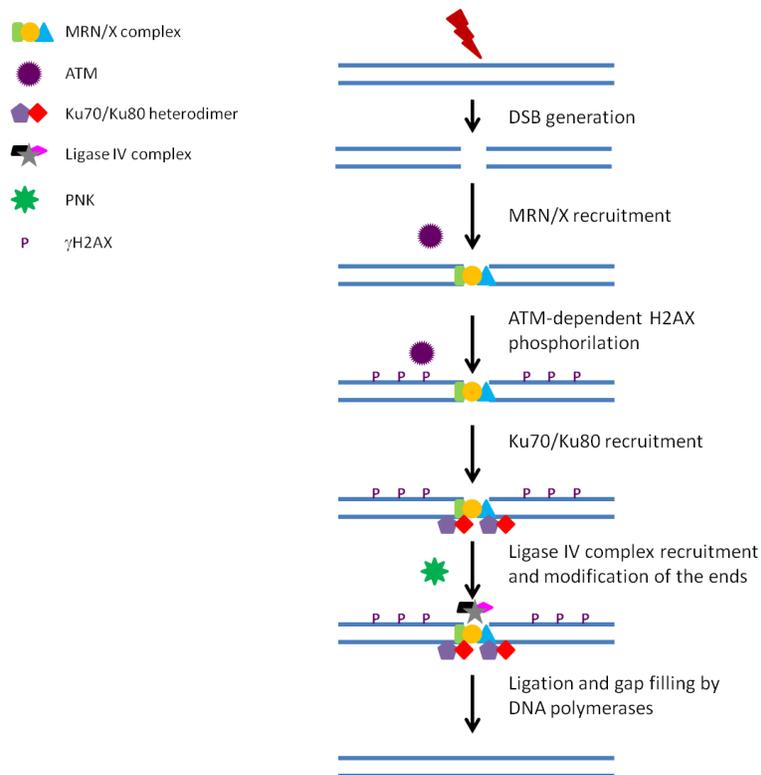


Figure 1.5. DSB repair by Non homologous end joining (NHEJ). In response to a DSB, the ATM kinase is activated and the MRN complex recruited to the site of lesion and the signalling starts. The ATM kinase phosphorylates the HA2X histone variant (γ H2AX). Ku is recruited to the DSB and in conjunction with the MRN complex maintains the ends of the DSB in close proximity. The Ligase IV complex is recruited and, if necessary, the 5' and 3' ends of the DSB are modified to generate ligatable ends. Ligation is carried out and the remaining gaps are filled by DNA polymerases.

1.4.2.2 Homologous recombination

HR is carried out mainly in late S or G2 phase of the cell cycle when a sister-chromatid is available. HR is a conservative mechanism that tends to restore the original sequence of the damaged DNA. For this purpose in HR the sister-chromatid is used as a template for synthesis of the new DNA at the site of the DSB. (Lord CJ *et al.* 2012). For initiation of the DSB repair through HR the MRN complex is rapidly recruited to the DSB and regulates 5'→3' resection at both sites of the DNA ends. The MRN complex interacts with CtIP (Mimitou EP *et al.* 2008), to initiate the resection of the DSB. Recent published data, from experiments in yeast (Mimitou EP *et al.* 2008), using exonuclease assays and ³²P labelled DNA molecules have proposed a model for the processing of DSBs. According to this model, the role of the MRX complex (yeast homolog of the MRN complex) and Sae2 (yeast homolog of CtIP) is to trigger the first cleavage reaction at the 5' ends of both strands using the helicase and flap endonuclease activities of the MRX complex and Sae2, respectively. This intermediate structure is rapidly processed by either the 5'→3' exonuclease activity of Exo1 or by Sgs1 (the yeast homolog of BLM). Sgs1 needs to act with a single strand specific nuclease to remove the 5' strand. The nuclease Dna2 seems a good candidate for this action (Mimitou EP *et al.* 2008).

After formation of the ssDNA tails at both sides of the DSB, they are coated by the Replication Protein A (RPA) (Lord CJ *et al.* 2012). RPA binding to ssDNA avoids formation of secondary structures within the ssDNA tails. Removal of secondary structures in the ssDNA is important for an adequate replacement of RPA by the RAD51 protein, and formation of the ssDNA-RAD51 filament. Rad51 is a member of the RecA-family of recombinases. *In vitro*, Rad51 mediates strand exchange between two homologous molecules of DNA in the presence of ATP. Rad51 plays key roles in homology searching and annealing of the ssDNA-RAD51 filament with the donor DNA molecule (Suwaki N *et al.* 2011), (Krejci *et al.* 2012). The formation of the ssDNA-RAD51 filament is an important regulatory point in the HR pathway. This step is facilitated by interaction of RAD51 with proteins called mediators. In humans the main mediator is BRCA2 (Liu J *et al.* 2011),

(Bugreev *et al.* 2007). On the other hand there are also proteins that act as “negative” regulators of the RAD51 functions like the RecQ helicases (Bugreev *et al.* 2007). Once the ssDNA-RAD51 filament is assembled, it searches and invades a dsDNA homologous molecule (first-strand invasion), generating a displacement loop (D-loop). Within the D-loop the invading strand primes the DNA synthesis. Once the DNA synthesis has initiated there are three different routes that the HR recombination pathway can follow (Figure 1.6). a) The invading strand can be unwound from its template after DNA synthesis and re-anneals with the resected end of the DSB. Thus the re-annealed strand provides the template for the synthesis of the second end of the DSB. This mechanism is called synthesis-dependent strand annealing (SDSA) and only produces non-crossover products (Figure 1.6 left) (Sugiyama T *et al.* 2006). b) The second tail of the DSB may anneal to the displaced DNA strand that is produced by DNA synthesis from the 3'-end of the first ssDNA (second strand annealing with the displaced strand). The annealing of the second strand pathways produces a double Holliday Junction (dHJ). This intermediate can be resolved to produce crossover or non-crossover products. Alternatively the dHJ can be dissolved to generate only non-crossover products (Figure 1.6 middle). c) In case the second ssDNA tail from the DSB is not available the D-loop may become a replication fork and the DNA synthesis progresses until the end of the chromosome. This process is known as break induced replication (BIR) and is not dependent on RAD51. The BIR mechanism is able to restore chromosome integrity, but it will cause loss of heterozygosity distal to the break point (Figure 1.6 right) (Sugiyama T *et al.* 2006), (Heyer WD *et al.* 2010).

SDSA seems to be the predominant pathway in somatic cells when both ends of the DSB are available (Heyer WD *et al.* 2010). This pathway avoids the generation of crossovers, and then reduces the possibility of chromosomal rearrangements. On the other hand, during meiosis the formation of dHJ and crossovers is important for recombination and chromosome segregation. Formation of dHJ and crossover products is not common in somatic cells. The BIR mechanism causes loss of heterozygosity distal to the DSB; this means that if there is a mutated allele in the DNA template it will be copied into the newly

synthesised DNA. Therefore the BIR pathway has potential risks for the cell that SDSA does not have (Heyer WD *et al.* 2010).

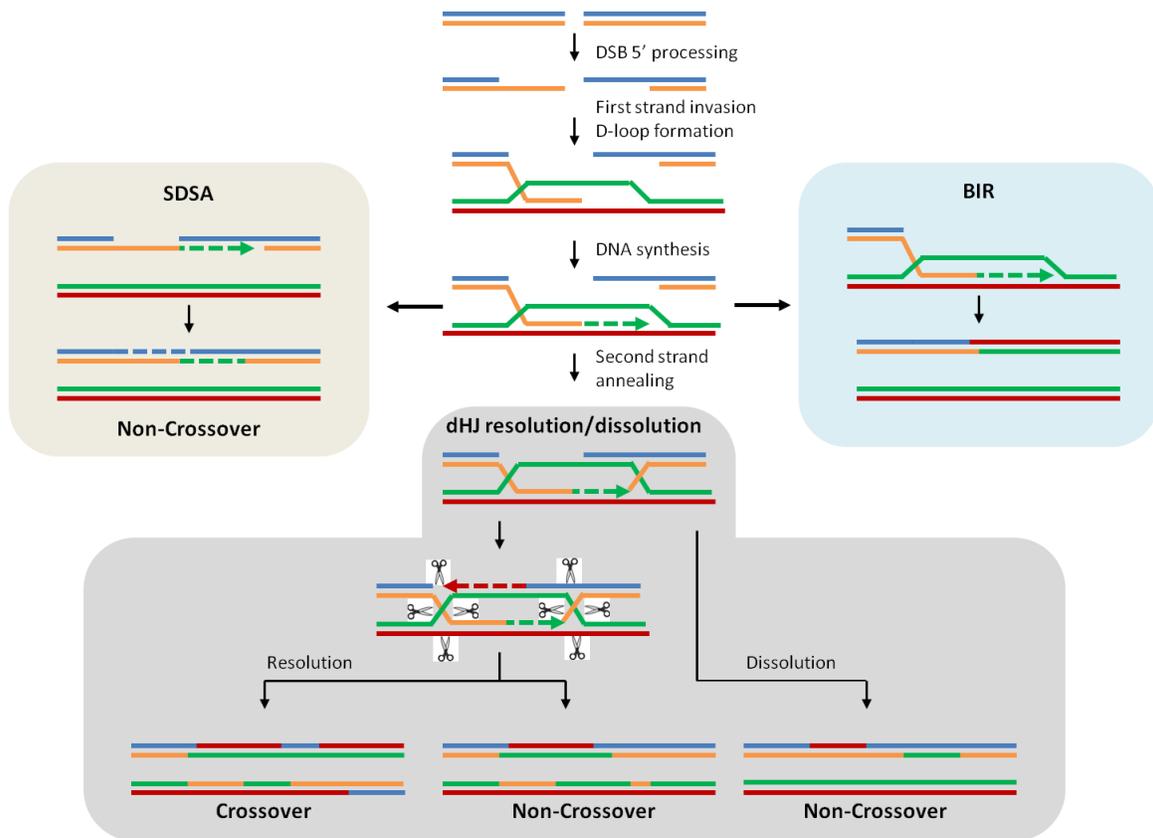


Figure 1.6. Models for DSB repair by HR Modified from (Sugiyama T *et al.* 2006). After recognition of the DSB, the first step is the processing of 5' ends to form long 3' ssDNA tails. When the homologous molecule is found, one ssDNA tail invades it (first strand invasion) and a displacement loop (D-loop) is formed. The invading strand primes DNA synthesis using the homologous molecules as a template. There are at least three possibilities to continue with the DSB repair (see text).

1.4.2.3 Microhomology mediated end joining

Microhomology mediated end joining (MMEJ) is a DSB repair pathway that uses microhomologous sequences (5-25bp) to align broken ends before ligation. This mechanism often results in deletions, insertions or translocations (McVey & Lee 2008).

MMEJ is distinct from NHEJ and HR pathways because it is independent of Ku and Rad52, although some authors have found that Rad52 is needed for some MMEJ events (Decottignies A 2007). The genes required in the MMEJ pathway have not been completely identified. Studies in fission yeast using mutants for genes used in NHEJ and HR repair have revealed that the Sae2, Tel1, Nej1 and Srs2 genes are important for the progression of the MMEJ pathway (Lee & Lee 2007). MMEJ is thought to occur preferentially, but not exclusively, at S and G2 phase (Decottignies A 2007). The mechanism and participants of the MMEJ are not completely known, but based on experiments in yeast some authors, (McVey & Lee 2008), have proposed a model for this pathway (Figure 1.7). The first step for the initiation of MMEJ is the resection of the 5' ends of the DSB in order to expose the regions of microhomology. DNA resection seems to be driven by the MRX complex, Exo1 and Sae2. The ssDNA generated can be coated by Rad51 to promote recombination, but Srs2 displaces Rad51 acting as an anti-recombination factor, leading DSB repair towards MMEJ. Annealing of the microhomology regions generate non-complementary 3' ssDNA tails. These flap structures must be removed before ligation, the candidate proteins for this function are Rad1-Rad10 structure specific endonuclease. After cleavage of the flap structures the DNA polymerases filled the remaining gaps. In yeast the DNA Pol4 activity is necessary for progression of the MMEJ pathway (McVey & Lee 2008). In human cells DNA pol β has been shown to efficiently fill the DNA gaps during MMEJ repair (Crespan E *et al.* 2012). When the joints have been annealed and the remaining gaps filled, ligation is carried out by ligases I and IV (McVey & Lee 2008). However, ligation in humans seems to be under the control of ligases I and III (Liang L *et al.* 2008). It is not fully understood how the MMEJ pathway is regulated. Ku seems to inhibit it and favours repair through NHEJ (Decottignies A 2007). Other genes that could be involved in MMEJ control are BRCA1 and the Chk2 protein kinase (Zhong Q *et al.* 2002). MMEJ is an area where more research is needed to understand the mechanics of the process and to know to what extent it contributes to the overall repair of DSB within the genome.

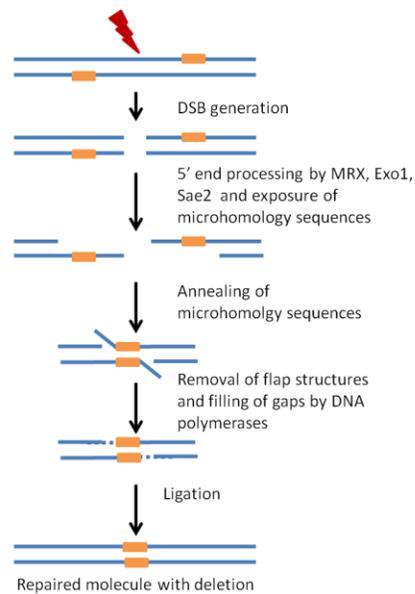


Figure 1.7 Microhomology mediated end joining (MMEJ). Microhomology mediated end joining (MMEJ) is a DSB repair pathway that uses microhomologous sequences (5-25bp) to align broken ends. In the first step the 5' ends of the DSB are processed presumably by MRX, Exo1 and Sae2 exposing the regions of microhomology. Microhomology regions are then annealed and the non-complementary 3' ssDNA tails form flap structures. The flap structures are removed and gaps filled by DNA polymerases. Ligation of the molecule finishes the process. As a result, the repaired DNA molecule has lost sequences at both sides of the DSB. Modified from (McVey & Lee 2008).

1.5 Telomere dysfunction

The disruption of the t-loop or associated telomeric chromatin structures by inactivating the genes encoding protein components required for telomere organisation and maintenance, by loss of the telomeric 3' overhang or by a extremely short tract of (TTAGGG)_n, can lead to telomere dysfunction or uncapping. Experimentally, depletion of shelterin components such as TRF2 or POT1 resulted in the colocalization of telomeres and DDR markers, like the phosphorylated form of the H2AX histone (γ -H2AX), in telomere-dysfunction induced foci (TIF) (Takai H *et al.* 2003). TIF formation is dependent on ATM or ATR protein kinases and the cellular arrest is mediated by p53, which occurs in response to telomere dysfunction, is thought to involve a DDR checkpoint (Takai H *et al.* 2003). It also has been observed that the disruption of TRF2 can induce NHEJ-dependent

leading to telomere:telomere chromosome fusion (Denchi EL *et al.* 2007), (Cesare AJ *et al.* 2009).

1.6 Telomere maintenance mechanisms

1.6.1 Telomerase

Telomerase was discovered in humans in 1997, as a multisubunit ribonucleoprotein enzyme. The catalytic subunit TERT (telomerase reverse transcriptase), shares seven domains with retroviral and retrotransposon retrotranscriptases (RT), but unlike other RTs, carries its own RNA template, the Telomerase RNA component (TERC). TERT and TERC form the core enzyme, because they are sufficient to provide enzymatic function *in vitro* (reviewed in Zvereva MI *et al.* 2010). In addition to these two components, telomerase contains other auxiliary components for its function *in vivo* which help to regulate its activity and to achieve telomere attachment in a cell cycle-dependent manner. Telomerase is able to elongate the chromosome 3'-end by adding TTAGGG repeats, whereas the complementary strand is completed by a kind of repair synthesis (Zvereva MI *et al.* 2010). One of the auxiliary factors important for telomerase activity is the dyskerin protein. Dyskerin is a pseudouridine synthase that catalyze the post-transcriptional conversion of uridine to pseudouridine in functional RNAs (Mitchell JR *et al.* 1999). In order to perform uridine modification, Dyskerine forms a complex with another three proteins (NHP2, NOP10 and GAR1) and a small RNA bearing the "hairpin-Hinge-hairpin-ACA" (H/ACA) motif. Mutations in the gene encoding dyskerin (*DKC1*) cause an X-linked form of dyskeratosis congenita (DC) (Mitchell JR *et al.* 1999). Cultured fibroblast or lymphoblast from DC patients showed reduced level of TERC, lower telomerase activity and shorter telomeres compared to matched controls (Mitchell JR *et al.* 1999). Studies have revealed that dyskerin and other H/ACA-motif RNA associated proteins are able to bind to the H/ACA motif in the TERC component of telomerase. Dyskerin binding to the TERT is necessary for proper maturation of the TERT precursor and stable accumulation of

the telomerase ribonucleoprotein (Mitchell JR *et al.* 2000), (Mitchell JR *et al.* 1999).

In humans the TERC component is widely expressed in most tissues, whereas TERT expression is strictly controlled and telomerase activity is primarily controlled by TERT expression (Zvereva MI *et al.* 2010). This is important because as mentioned previously, telomere shortening helps to regulate cell division potential and acts as a tumor suppressor mechanism. Telomerase activity is found in embryonic stem cells and some proliferative reproductive cells, and its activity can be detected in testes and ovary, but not in mature spermatozoa or oocytes (Zvereva MI *et al.* 2010). Telomerase activity can also be detected in some progenitor/tissue specific stem cells. In somatic tissues telomerase activity usually is not present. The regulation of the expression of telomerase has been extensively studied. The promoter region of the *TERT* gene contains binding sites for several transcription factors, and some of them can be related to the activation/regulation of telomerase in cancer cells. For example some proteins are c-Myc, the estrogen receptor, the oncoprotein E6 from HPV16 (Kyo S *et al.* 2008).

Recent studies have showed that the *TERT* promoter is also regulated by epigenetic factors. The inhibition of classes I and II of histone deacetylases (HDACs) by trichostatin A is able to induce expression of *TERT* in normal and telomerase negative cells lines. This is supported by the finding of hyperacetylation of the core histones at the *TERT* promoter in several cell lines that express telomerase activity. The methylation status of the cytosines at the promoter region of *TERT* has been largely studied, but contradictory results have been found. Some authors proposed that the methylation of the *TERT* promoter cause silencing of the gene (Liu L *et al.* 2004), (Lopatina NG *et al.* 2003), while some others suggest that there is no a general correlation between cytosine methylation at the *TERT* promoter and its expression (Dessain SK *et al.* 2000), (Devereux TR *et al.* 1999). Interestingly, some authors have reported that hypermethylation of the CpG sites at the *TERT* promoter causes overexpression of the gene in some cancer cell lines, this is contrary to what is commonly observed as hypomethylation of the cytosines at the *TERT* promoter allows binding of the repressor factor CTCF (Renaud S *et al.* 2007). In addition to

the histone and cytosine modifications, it is thought that integration of viruses and chromosomal rearrangements can trigger *TERT* transcription (Zhu J *et al.* 2010). These studies reveal the complexity of *TERT* locus regulation.

1.6.1.1 Diseases associated with telomerase

Dyskeratosis congenita (DC) is a clinical condition associated with mutations in telomerase components (Savage SA *et al.* 2010). Clinical features of DC encompass nail dystrophy, abnormal pigmentation, eye and dental abnormalities, pulmonary fibrosis and liver disease. In addition, patients with DC are at increased risk of progressive bone marrow failure (BMF), myelodysplastic syndrome or acute myeloid leukaemia (AML), and solid tumors. Approximately 50% of DC cases are caused by mutations in one of the genes related to telomerase. In 2010 Savage *et al.*, published a review summarizing molecular, clinical and epidemiological features of the disease. Table 1.1 (taken from Savage SA *et al.* 2010).

In addition to classical DC, two severe forms of DC exist. Hoyeraal-Hreidarsson syndrome is characterized by DC symptoms plus cerebellar hypoplasia and developmental delay, and Revesz syndrome which includes DC features plus retinopathy and intracranial calcifications. *DKC1* and *TINF2* have been found to be mutated in Hoyeraal-Hreidarsson patients and *TINF2* in Revesz patients (Savage SA *et al.* 2010).

The *TERT* gene (located in chromosome 5) has also been found to be deleted in patients with Cri du chat syndrome, which is caused by a deletion of variable size in the short arm of chromosome 5. Zhang *et al.*, noticed that a group of patients with Cri Du Chat syndrome had shorter telomeres and a high proportion of telomere fusions in cultured fibroblast, compared with age-matched controls, and they propose that *TERT* haploinsufficiency could contribute to the phenotype of Cri du chat patients (Zhang A *et al.* 2003).

Table 1.1 Diseases associated with telomere biology dysfunctions and genes involved. The table describes clinical diseases associated to mutations in components of the enzyme telomerase. The numbers represent the percentage of patients with mutations in the corresponding gene, except when it is specified as families. DC dyskeratosis congenita, SAA severe aplastic anemia, IPF idiopathic pulmonary fibrosis, AML acute myelogenous leukemia, XLR X-linked recessive, AD autosomal dominant, AR autosomal recessive, NR none reported.

Gene, protein	Chromosomal locus	Inheritance	Approximate percent of patient with mutations				
			DC	Isolated SAA	Isolated IPF	Isolated fibrotic liver disease	Isolated AML
<i>DKC1</i> , dyskerin	Xq28	XLR	17-36	NR	NR	NR	NR
<i>TINF2</i> , TIN2	14q11.2	AD	11-24	5	NR	NR	NR
<i>TERC</i> , TERC	3q26.3	AD	6-10	1-5	1-3	3 Families	NR
<i>TERT</i> , TERT	5P15.53	AD, AR	1-7	3-4	7-12	2 Families	2
<i>NOP10</i> , NOP10	15q14-q15	AR	<1	NR	NR	NR	NR
<i>NHP2</i> , NHP2	5q35.5	AR	<1	NR	NR	NR	NR

1.6.2 The ALT mechanism

The ALT pathway is a telomerase-independent telomere length maintenance mechanism. It is present in ~15% of human cancers with a higher prevalence in tumors of mesenchymal origin, but it also can be detected in a small proportion of common cancers, such as breast cancer. Among the tumors with high prevalence of ALT activity are glioblastoma multiforme, osteosarcoma. Clinical studies have suggested that ALT positive tumors can be associated with poor prognosis (Cesare AJ *et al.* 2010).

1.6.2.1 Hallmarks of ALT positive cells

Several phenotypic features are associated with the ALT pathway. These features make it possible to distinguish ALT cells from telomerase-positive or normal cells:

a) Extrachromosomal telomeric DNA. This can be found in different forms, the most abundant being double stranded telomeric circles (t-circles) that can be the result of the resolution of t-loops by recombination enzymes. It has been proposed that instead of being involved in the ALT mechanism. The t-circles could represent the final product of a trimming mechanism in ALT cells, due to the extensive lengthening of telomeres. This trimming mechanism is designed to control telomere length, something analogous to the rapid deletion mechanism observed in yeast (Cesare AJ *et al.* 2010).

Partially single stranded circles represent another form of extrachromosomal DNA, which can be called C-circles or G-circles according to the strand that is continuous (Nabetani A *et al.* 2009). Also linear dsDNA and high molecular weight "t-complex" DNA thought to contain branched structures, can be found (Nabetani A *et al.* 2009). Recently Henson *et al.*, have shown that C-circles are 100-fold more abundant than G-circles and that the disruption of the ALT pathway leads to a significant decrease of C-circles. Even if the origin of the C-circles is not very clear, they have been found in blood samples from patients

with ALT positive tumors which make them a potential diagnostic tool, as well as an approach to monitor ALT activity in tumors. These findings prompted the suggestion that C-circles are the most useful marker for ALT activity (Henson JD *et al.* 2009).

b) Promyelocytic leukaemia bodies (PML) associated with ALT (APB). PML bodies are structures normally found in cells in interphase, their main components are the PML and the SP100 proteins. Under normal conditions no DNA is found in the PML bodies (Yeager TR *et al.* 1999). APBs are defined by the presence of telomeric DNA and associated shelterin proteins, as well as various proteins involved in DNA replication, recombination and repair, such as helicases (BLM and WRN), RPA, the MRN complex and RAD52 (Yeager TR *et al.* 1999), (Nittis T *et al.* 2008). The role of APBs in ALT cells is still under debate. One idea is that they are the location of ALT activity. Among the findings that support this hypothesis are, that the MRN complex is necessary for APB formation (Jiang WQ *et al.* 2005). The interference with ALT activity results in a decrease of APBs, and it has been observed that APBs are more abundant in G2 when recombination is more active (Cesare AJ *et al.* 2010). More evidence supporting that APBs are involved in the ALT pathway was provided by Draskovic I, *et al.*, they show that telomere-telomere interactions occur within APBs (Draskovic I *et al.* 2009). However, APBs have been observed in telomerase-positive cells lines after the overexpression of *TERT* sequestering extrachromosomal DNA. No other evidence of ALT activity was found (Pickett HA *et al.* 2009). Furthermore, in telomerase-negative liposarcomas the presence of ALT markers such as MS32 instability, and complex telomere mutations have been observed in the absence of APBs (Jeyapalan JN *et al.* 2008). These observations and the identification of cell lines using the ALT mechanism without the presence of APBs indicate that they are not required for ALT activity (Cerone MA *et al.* 2005).

c) Heterogeneous telomere length. The mean length of telomeres in telomerase-positive cell lines is <10kb, while in ALT+ cell lines telomeres can range from <2 kb to >50 kb. ALT dynamics are complex, with both rapid elongation and rapid deletion events. In addition

such events may generate linear or circular extrachromosomal telomeric DNA (Fasching CL *et al.* 2007).

d) Minisatellite MS32 instability. The minisatellite MS32 has been proved to be highly unstable in ALT cell lines and tumors when compared to normal or telomerase positive cell lines (Jeyapalan JN *et al.* 2005). Other minisatellites have been investigated but none of them showed the levels of instability as MS32. The underlying mechanism and the relation to the ALT pathway are not fully understood (Jeyapalan JN *et al.* 2005).

e) Telomere sister chromatid exchange (TSCE). ALT cells exhibit a higher rate of T-SCE compared with telomerase positive cell lines, it is important to notice that the increase in SCE is not observed in the rest of the genome. Chromosome-orientation fluorescence *in situ* hybridisation (CO-FISH) can be used to detect T-SCE events (Londono-Vallejo JA *et al.* 2004). This method is based on allowing the cells to replicate their DNA once in the presence of a bromo-substituted nucleotide (Bromodeoxyuridine or Bromodeoxycytosine). After replication the daughter strand will contain one parental strand and a newly-synthesised strand containing the bromo-substituted nucleotide. Metaphases are prepared and exposed to UV light for degradation of the bromo-substituted nucleotide. By using C-rich or G-rich probes CO-FISH can distinguish between the leading and lagging strand (Bailey SM *et al.* 2001). Because of the repetitive nature of telomeric DNA exchange between telomeres are often unequal. This unequal exchange results in one telomere growing longer at the expenses of the sister chromatid. These events contribute to the heterogeneous telomere length observed in ALT+ cells (Londono-Vallejo JA *et al.* 2004).

1.6.2.2 Homologous Recombination and ALT models

The first evidence of a telomerase-independent TMM was observed in cells of the budding yeast *S. cerevisiae* inactivated for telomerase (*tlc1Δ*). These cells are able to survive by maintaining telomeres using a HR-dependent mechanism (Meringele L *et al.* 2004). Two types of survivors have been reported in yeast *tlc1Δ*, Type I and Type II. Both types are dependent on function of the recombination protein Rad52 (Chen *et al.* 2001). Type I survivors, are additionally dependent on Rad51. Type I survivors amplify their subtelomeric Y' elements adjacent to the telomeres. During this process, a short telomere is processed as a DSB with just one end available, to generate a long 3' ssDNA tail. The ssDNA tail invades the Y' element of another telomere and then DNA synthesis progresses until the end of the chromosome (Chen *et al.* 2001). On the other hand, Type II survivors amplify the TG₁₋₃ telomeric repeats and this is dependent of the MRX complex (Mre11-Rad50-Xrs2), which is the counterpart of the MRN complex in humans (Chen *et al.* 2001). It has been shown that RAD50 and NBS1 play an important role in the ALT pathway (Jiang WQ *et al.* 2005). In addition, type II survivors in *S. cerevisiae* show extensive elongation of the TG₁₋₃ telomeric repeats up to 10kb and heterogeneity in the length of telomeres is observed (McEachern MJ *et al.* 2006). These findings suggest that human ALT cells are analogous to the Type II survivors of yeast and showed a connection between the ALT pathway and the recombination/repair machinery (Nabetani A *et al.* 2009). Another finding that supports the HR-dependent nature of the ALT pathway is the fact that a sequence inserted in one telomere can be copied to other telomeres in ALT cells but not in telomerase-positive cells (Reddel RR *et al.* 2001). In order to explain how telomeres can be extended by a HR-like mechanism, different models have been proposed (Figure 1.8 modified from Cesare AJ *et al.* 2010):

a) Unequal Telomere Sister Chromatid Exchange (TSCE). The mechanism for TSCE is not fully understood yet, but it has been proved that telomeric DNA contains gaps and nicks in ALT+ cells (Nabetani A *et al.* 2009). The model proposes that the gaps and nicks cause

stalling of replication forks, and that TSCE emerges as a way to re-initiate DNA replication (Cesare AJ *et al.* 2010). This model was proposed because SCE elsewhere in the genome can result from repair, through recombination, of stalled replication forks (Wilson DM 3rd *et al.* 2007). The outcome of the unequal TSCE is a very short telomere with a proliferative disadvantage and a chromosome with an extended telomere which could give proliferative advantage to daughter cells (Figure 1.8A) (Cesare AJ *et al.* 2010).

b) Homologous recombination-dependent replication model. Another theory is that the telomere could use another telomere as a template. This model is supported by the observation of a sequence copied from one telomere to another in ALT cell lines, but not in telomerase positive cells (Dunham MA *et al.* 2000). In addition, it is known that at the beginning of the telomeres there are some telomere variant repeats interspersed with the TTAGGG repeats. The pattern of TTAGGG and variant repeats can be obtained by telomere mapping and it is characteristic for each chromosome (Varley H *et al.* 2002). Telomere mapping in ALT+ cells has shown changes in the pattern of interspersion of TTAGGG and variant repeats at a specific telomere. Some of these mutations generate maps completely unrelated to the progenitor map. Suggesting that they have been copied from another telomere or extrachromosomal telomeric DNA. These mutations have been called complex mutations and they are not observed in normal or telomerase-positive cell lines (Varley H *et al.* 2002).

In the HR-dependent model the potential template molecules include t-loops, linear or circular extrachromosomal DNA and the sister chromatid, or a non-homologous telomere. Figure 1.8 B-E (Cesare AJ *et al.* 2010).

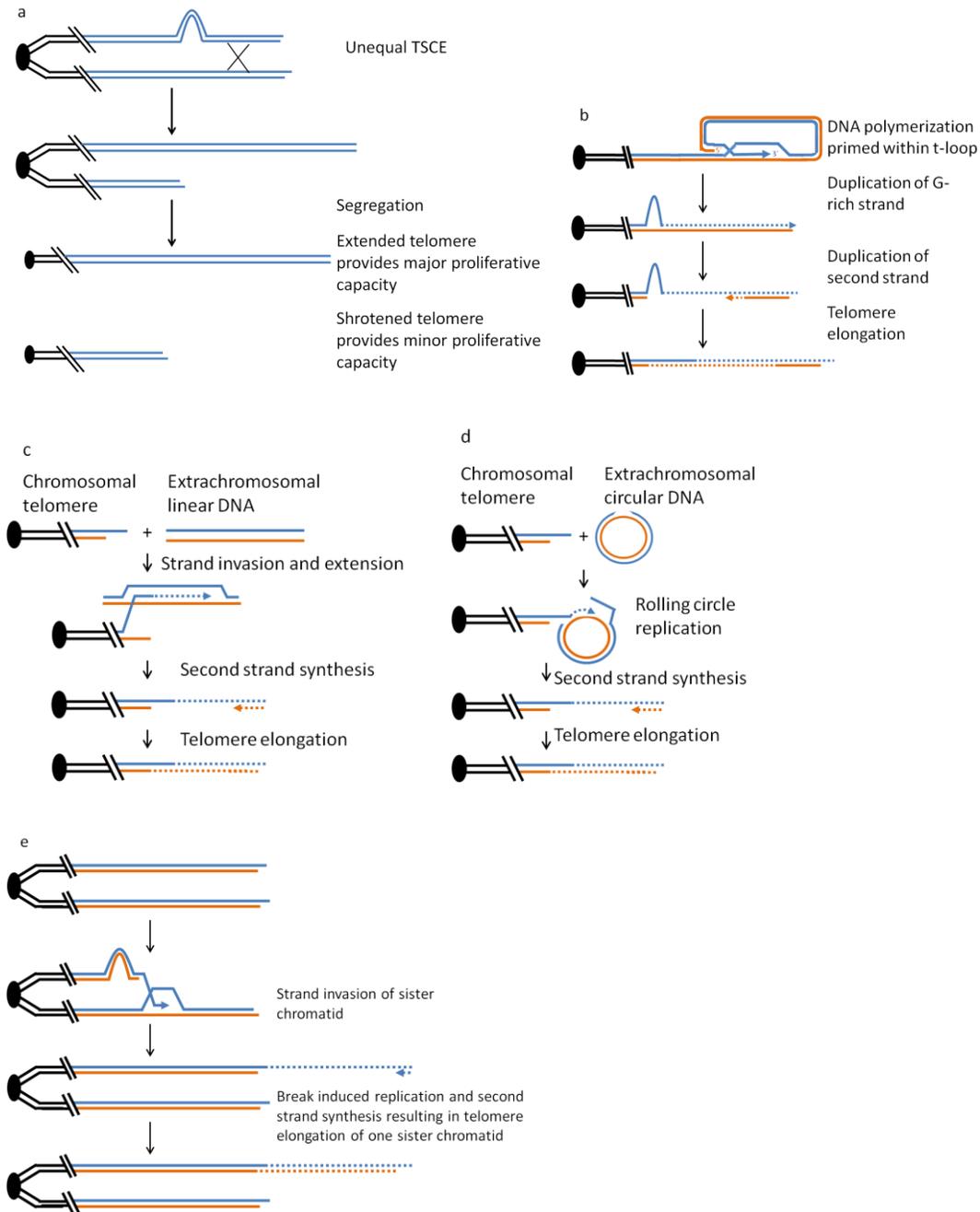


Figure 1.8 Proposed models for telomere elongation in ALT cells. (A) Unequal TSCE. The result of this mechanism is one chromosome with a very extended telomere and the other with a very short telomere. After segregation the longer telomeres provide a proliferation advantage to the cells. (B) Intramolecular telomere amplification this is primed by the 3' overhang in the t-loop. (C, D) Use of extratelomeric, linear or circular, DNA as template for telomere extension. (E) Strand invasion between sister chromatides resulting in telomere extension. Alternatively this also can happen between non homologous telomeres. Modified from (Cesare AJ et al. 2010).

In yeast the HR repair pathway called BIR has been described. This pathway involves a type of homologous recombination repair as it starts with the 5' to 3' degradation of the broken chromosome end to generate a long 3' single stranded tail. This tail then initiates strand invasion and pairing with a homologous DNA molecule. Usually the 5' resection in a DSB occurs at both ends of the break in a way that the generated 3' tails can initiate strand invasion and pairing with a homologous sequence, leading to a conservative repair process (Llorente B *et al.* 2008). In BIR, only one end of the DSB undergoes homology-dependent strand invasion, and then the 3' tail primes DNA replication. DNA synthesis proceeds for many kilobases to the end of the chromosome and this process seems to involve leading and lagging strand synthesis, while other HR pathways need only leading strand synthesis (Llorente B *et al.* 2008). Some authors have proposed that the BIR and ALT mechanism work in the same way (Reddel *et al.* 1997). This has led to the study of genes involved in DNA recombination/repair in the ALT pathway. Among the most studied genes are *Sgs1*, the yeast helicase that shares homology to the Werner Syndrome (*WRN*) and Bloom Syndrome (*BLM*) genes. *Sgs1* seems to be important for the generation of survivors in the absence of telomerase and RAD51 (Johnson FB *et al.* 2001). *WRN* and *BLM* belong to the family of RecQ helicases in humans. There is evidence suggesting that they play important roles on telomere maintenance (Bohr VA 2008).

1.7 RECQ Helicases

The RecQ protein family encodes a highly conserved group of DNA helicases which participate in several different metabolic processes, such as DNA replication, repair, recombination and transcription. In *E. coli*, the fission and budding yeasts only one RecQ homolog has been identified, whereas in humans there are five. The identified members of the RecQ gene family in humans are *WRN*, *BLM*, *RECQ1*, *RECQ4* and *RECQ5* (Bohr VA 2008). In humans WS is characterized by genomic instability accompanied by premature aging and a higher risk of develop tumors especially sarcomas and meningiomas. It has

been shown that cells from WS patients display accelerated telomere loss (Crabbe L *et al.* 2007). Nevertheless, a study published by Baird DM *et al.*, in 2004, showed that the telomere erosion rates at the XpYp telomere from clones of the W-V cell line (WRN^{-/-}) were not different from those in normal fibroblast when measured by STELA (Single telomere amplification). This suggests that the accelerated senescence observed in WS cells is not related to an accelerated telomere erosion (Baird DM *et al.* 2004).

BS is characterized by short stature and also a high risk of cancer (Bohr VA 2008). It is known that BLM can interact with TRF1 and TRF2 and might play a crucial role in cells using the ALT pathway (Lillard-Wetherell K *et al.* 2004). As mentioned before mutations in the WRN and BLM helicases are associated with WS and BS syndrome respectively, while mutations in the *RECQ4* gene are associated with Rothmund Thompson (RTS), RAPADILINO and Baller Gerold (BGS) syndromes (Bohr VA 2008). To date no phenotypes have been associated with mutations in *RECQ1* or *RECQ5*. That WS, BS and RTS patients all show a high predisposition to cancer reveals the important role of these enzymes in maintenance of genome stability (Bohr VA 2008).

1.7.1 WRN and BLM helicases

The WRN and BLM helicases unwind DNA in a 3' to 5' polarity; the favourite substrates for these enzymes are HR intermediates like Holiday Junctions (HJ) and G-quadruplex structures in telomeric DNA (Bohr VA 2008). Unlike other members of the RecQ family, WRN and BLM helicases possess DNA-dependent ATPase activity and ssDNA-annealing activity (Wu L *et al.* 2006). In addition, WRN is unique among the RecQ helicase family because it has an exonuclease activity in a 3' to 5' direction. The similarity in function between the WRN and BLM helicases raises the question about functional co-operation between them. Experimental data have demonstrated colocalization of WRN and BLM in human cell lines, and coimmunoprecipitation analysis has showed physical interaction

between them (von Kobbe C *et al.* 2002). However, functional analysis has revealed lack of synergistic unwinding activity by WRN and BLM. It has been demonstrated that BLM is able to inhibit WRN exonuclease activity (von Kobbe C *et al.* 2002). Several proteins have been shown to interact with WRN and BLM, most of them playing roles in DNA repair or recombination. In addition, both helicases are known to interact with members of the shelterin complex POT1, TRF1 and TRF2 (Lillard-Wetherell K *et al.* 2004, Rossi ML *et al.* 2010). The details of these interactions will be discussed later.

1.7.2 WRN and BLM in telomere maintenance

The phenotype of WRN and BLM deficient cells include telomere dysfunction, pointing to some role in telomere maintenance. In *BLM* deficient cells there is an increase in telomere associations between homologous chromosomes, whereas cells lacking *WRN* shown increased telomere loss (Crabbe L *et al.* 2007). Ectopic expression of telomerase can rescue primary fibroblasts from BLM and WRN donors from premature senescence (Wyllie FS *et al.* 2000). The KO mouse model null for *WRN* and telomerase (double KO) appears normal, however in later generations, when telomeres are shorter, a cellular and clinical phenotype closely resembling WS patients is observed (Chang S *et al.* 2004). On the other hand, the mouse null for *BLM* is not viable, which means that *BLM* is indispensable for mouse development (Luo G *et al.* 2000). However the generation of a hypomorphic allele allowed the production of a mouse model with a phenotype very similar to that observed in BS patients. These mice displayed high rate of SCE and elevated predisposition to cancer (Luo G *et al.* 2000). A mouse model null for both *BLM* and *WRN*, generated in a telomerase-negative animal, showed a more severe phenotype caused by a synergistic, more than additive effect. Cells from this model showed increase telomere loss and chromosomal end fusions compared to wild type (Du *et al.* 2004).

Colocalization and coimmunoprecipitation studies have revealed affinity of TRF2 for WRN

via the RQC conserved domain present in members of the RecQ family (Opresko PL *et al.* 2004). *In vitro* analysis using ELISA showed that TRF2 is also able to interact with BLM via the same domain. Analysis using telomeric substrates *in vitro* showed that TRF2 is able to stimulate the unwinding activity of both WRN and BLM proteins, but the nuclease activity of WRN is not enhanced in the presence of TRF2 (Opresko PL *et al.* 2002). The ability of BLM and WRN to unwind substrates containing telomeric repeats previously coated with TRF2 suggests that both enzymes can be recruited by TRF2 to contribute to the resolution of secondary structures present at telomeres during replication or repair (Opresko PL *et al.* 2002).

It is known whether RPA helps BLM and WRN to unwind DNA molecules by binding ssDNA (Nimonkar AV *et al.* 2011). POT1 binds to ssDNA but has specificity for telomere sequences. *In vitro* studies using recombinant full-length POT1 showed that it can stimulate the helicase activity, but not the exonuclease activity of WRN on substrates containing telomeric repeats resembling a t-loop structure. In a similar way, the presence of recombinant POT1 increased the unwinding activity of BLM on the same structures (Opresko PL *et al.* 2005). Analysis using ELISA proved that POT1 physically interacts with BLM and WRN, but the interaction with WRN is slightly greater. Based on these findings, it has been proposed that the interaction of POT1 with BLM and WRN facilitates the coating by POT1 of the ssDNA telomeric DNA that is generated by the unwinding activity of WRN so avoiding reannealing of the duplex DNA (Opresko PL *et al.* 2005).

Consistent with the role of WRN in telomere maintenance, the expression of a dominant negative helicase-deficient form of WRN in HeLa cells leads to an increase in the number of chromosomes showing loss of telomere signal on a single chromatid, (Sister Telomere Loss, STL) (Crabbe *et al.* 2004). The proportion of cells exhibiting STL increased after inhibition of the TERT subunit of telomerase. In addition, in cells from WS patients the expression of wild type WRN or an exonuclease-deficient mutant reduce the number of STL, however the expression of a helicase-deficient mutant did not change the number of

STL observed. Chromatin immunoprecipitation (ChIP) analysis have demonstrated that association of WRN with telomeres occurs specifically in S phase of the cell cycle, pointing to a function of WRN during DNA replication. Using Chromosome Orientated FISH (CO-FISH), it was shown that the lagging strand synthesis was more affected in HeLa cells than after the expression of a helicase-deficient form of WRN. A possible explanation is that during lagging strand synthesis, G-quadruplex structures can be formed on the G-rich strand causing stalling of replication fork. WRN has the ability to resolve these structures, therefore in the absence of WRN, replication stalls with consequent loss of telomere signal (Crabbe *et al.* 2004). It also has been observed that a higher percentage of cells from WS patients display DDR signals, for example ATM phosphorylation on serine 1981 (Crabbe *et al.* 2004). Expression of wild type *WRN* or *TERT* reduces the percentage of cells showing DDR. In addition, the number of anaphase bridges observed in WS cells after suppression of the Rb and p53 pathways significantly decrease when *TERT* is expressed. These results highlight the importance of WRN in maintaining telomere stability and genomic integrity (Crabbe *et al.* 2004).

1.8 The EXO1 exonuclease

The human exonuclease 1 gene (*EXO1*) is located at 1q42-q43. It contains 1 untranslated exon followed by 13 coding exons (Tran PT *et al.* 2004). *EXO1* is a member of the RAD2 family of structure-specific nucleases which possess 5'→3' exonuclease and 5'-flap endonuclease activities *in vitro* (Tran PT *et al.* 2004). Two splice variants *EXO1a* and *EXO1b* have been identified and the transcripts give rise to proteins of 803 and 846 amino acids, with predicted molecular weights of 89 and 94 kDa, respectively. Both isoforms contain the exonucleolytic domain and the ratio of the different transcripts has been reported to be 1:6 for *EXO1a*:*EXO1b* (Tran PT *et al.* 2004). So far, no biological or biochemical differences between the two proteins have been reported (Knudsen NO *et al.* 2007).

The *EXO1* gene encompasses 41.5kb of genomic DNA and it has 666 SNPs, as reported by the NCBI (www.ncbi.nlm.nih.gov/gene/9156). Among these 27, are located in the coding sequence (CDS); 6 are synonymous, 18 missense and 3 frame shift. From the 18 missense, the Ser610Gly and the Gly759Glu were found more frequently in patients with Hereditary Polyposis Colorectal Cancer (HNPCC) by Wu, *et al* in 2001, (Wu Y *et al.* 2001). However in 2002 Sun *et al*, performed functional analysis of the potential mutations described by Wu, and they found that the products of these variants had enzymatic activity. As it is known that *EXO1* interacts with the components of the DNA MMR system *MLH1*, *MSH2*, and *MSH3*, Sun *et al*, tested the ability of these variants to interact with *MLH1* and *MSH2* using co-immunoprecipitation assays. The amount of *MLH1* that precipitated with *EXO1* mutants was similar to that observed for the wild type, and for *MSH2* they observed that the variant Gly759Glu pulled-down approximately one third of the amount that was observed with the wild type. Interestingly this mutation is located in the domain for interaction with *MSH2*, indicating that this residue is important for the interaction between MMR proteins (Sun X *et al.* 2002)

The RAD2 family of nucleases is conserved from bacteriophage to humans and consists of a group of nucleases that participate in several biological pathways such as DNA repair and recombination. The members of this family possess a conserved nuclease core, which is contained in the "N" (N-terminal) and "I" (internal) domains (Lee Bi BI *et al.* 2002). The eukaryotic proteins are divided in three groups according to amino acid sequence identity, position of the nuclease domains and substrate specificity (Lee Bi BI *et al.* 2002). The RAD2 class I consists of the *Xeroderma Pigmentosum* complementation group G (XPG)-like proteins that participate in the NER. Mutations in the *XPG* gene are associated with the disorder *Xeroderma Pigmentosum*, characterized by hypersensitivity to sunlight and an increased risk of developing skin cancer (de Boer J *et al.* 2000). The RAD2 class II family is composed of the *FEN 1*-like proteins. These enzymes have a flap endonuclease activity for branched DNA structures produced by polymerase strand displacement or as intermediates during DNA recombination. RAD2 class III consists of the Exo1-like enzymes

found in yeast, fly and mammals. Exo1 was first identified in *Schizosaccharomyces pombe* as a meiotically induced 5'→3' exonuclease (Tran PT *et al.* 2004). Human EXO1 shares 27% amino acid identity with yeast Exo1 (Tishkoff DX *et al.* 1998), and has been shown to have 5' to 3' exonuclease activity on ssDNA and dsDNA as well as flap structure-specific endonuclease activity, which is dependent on Mg²⁺ (Lee BI *et al.* 1999). It is known that Exo1 can degrade blunt ends to produce 3' ssDNA tails, and also from nicks (Tran PT *et al.* 2004).

It has been demonstrated that EXO1 is phosphosrylated in basal conditions at 9 amino acids (Ser₃₇₆, Ser₄₂₂, Ser₅₉₈, Ser₆₂₃, Ser₆₃₉, Ser₆₆₀, Ser₆₇₄, Ser₇₄₆ and Thr₅₈₁) and it can be phosphorylated at three additional residues in the presence of inhibitors of DNA replication in an ATR dependent manner (Ser₄₂₄, Ser₆₂₁ and Ser₇₁₄) (El-Shemerly M *et al.* 2007). The final result of this phosphorylation is the degradation of EXO1 by the proteasome pathway. The substitution of just one or two of these amino acids to alanine does not result in the inhibition of the protein degradation. However, the triple mutant shows increased stability compared to the wild type. This confirms the importance of the three residues in the degradation of EXO1 (El-Shemerly M *et al.* 2007).

EXO1 has domains for interaction with three members of the DNA MMR system, including MLH1, MSH2 and MSH3. The EXO1 protein domains are displayed in Figure 1.9, and table 1.2.

Table 1.2 Position, length and function of the domains of the human EXO1 protein (Uniprot www.uniprot.org/uniprot/Q9UQ84).

Position aa	Length	Region
1-99	99	N-domain (nuclease)
129-387	259	Interaction with MSH3
138-229	92	I-domain (nuclease)
388-490	103	Interaction with MLH1
418-421	4	Nuclear localization signal
600-846	247	Interaction with MSH2
787-846	60	Interaction with MLH1

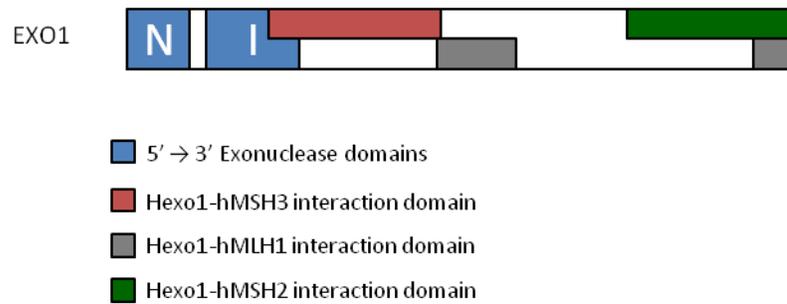


Figure 1.9. EXO1 functional domains. Modified from (Tran PT *et al.* 2004).

1.8.1 EXO1 function in DSB repair by Homologous Recombination

Experiments in yeast have shown that deletion of the *Exo1* gene resulted in altered recombination frequency. In addition extracts from *Exo1*-deficient *S. cerevisiae* strains were less effective at recombining linear DNA molecules possessing overlapping homologous ends (Wilson DM 3rd *et al.* 1998). Furthermore, *Exo1* mutants were found to display a 6-fold reduction in the rate of meiotic recombination between non-tandem duplications. These findings and the fact that *Exo1* encodes a 5'→3' exonuclease that generates single stranded complementary tails from ends or nicks, which promote joint molecule formation, suggest that *Exo1* homologs function as 5'→3' exonucleases in the homologous recombination and mismatch repair processes (Wilson DM 3rd *et al.* 1998). However there is evidence that *Exo1* could have overlapping roles with some other nuclease in the processing of DSBs (Tran PT *et al.* 2004). Experiments using exonuclease assays, *in vitro*, with radio-labelled DNA molecules have proposed a model for the processing of DSBs. In this model the role of the MRX complex and Sae2 is to bind to the DSB and trigger the first cleavage reaction at the 5' ends of the DSB using the helicase and flap endonuclease activities of the MRX complex and Sae2 (yeast homolog of human CtIP discussed below) respectively (Mimitou EP *et al.* 2008). Then this intermediate is rapidly processed by either the 5'→3' exonuclease activity of *Exo1* or by the coordinated activity of the yeast homolog of BLM, (*Sgs1*) in conjunction with a single strand specific nuclease

to remove the 5' strand, such as the nuclease Dna2 (Mimitou EP *et al.* 2008). The authors observed that in wild type yeast strains the velocity at which DSBs are processed is approximately 4.4 kb/hr. This rapid processing was affected upon deletion of Sgs1 or Dna2. Under this scenario Exo1 seems to be responsible for the 5' processing and the velocity was slowed down to approximately 1 kb/hr (the velocity of DNA processing is estimated by the presence or absence of DNA regions next to the DSB after a determined period of time). However, the most severe effect in 5' processing was observed on double mutants, either *Sgs1 Exo1* or *Dna2 Exo1*. Under these circumstances the 5' processing was very slow and is thought to be carried out by the MRX complex and Sae2. In addition to the velocity of processing they observed that the efficiency was also affected because not all the cells were able to initiate the 5' processing (Zhu Z *et al.* 2008).

In yeast there are different pathways for DSB repair through HR. One of these is called break induced replication (BIR) pathway. As in other pathways, the BIR mechanism starts with the generation of a long 3' overhang, formed by degradation of the 5' end. This end initiates the strand invasion, but, unlike other mechanisms in the BIR pathway just one end undergoes strand invasion, and the DNA synthesis progresses until the end of the chromosome (Llorente B *et al.* 2008). In order to accomplish full DNA replication BIR requires leading and lagging strand synthesis, this is also different from other HR pathways (Llorente B *et al.* 2008). One of the candidate exonucleases for the processing of the 5' end and progression of the BIR mechanism is Exo1. This is supported by the fact that some BIR events are dependent on Exo1 (McEachern MJ *et al.* 2006), and that deletion of Exo1 delays the emergence of type I and Type II survivors of yeast strains deficient in telomerase (Maringele L *et al.* 2004). Nevertheless, recent experiments in yeast have proposed that Exo1 and Sgs1 can inhibit the progression of the BIR mechanism. These results were obtained with an assay designed to monitor the progression of BIR. This assay consists in introducing a plasmid into the cell. Then the linearization of the plasmid expose a region with homology to chromosome III at one end and yeast telomeric repeats at the other end. The observation was an increase in the rate of BIR events after

deletion of *Exo1* and *Sgs1*. However an important consideration is that the plasmid was completely degraded after linearization in the presence of *Exo1* and *Sgs1*. The degradation of the plasmid can be explained by the extensive processing by *Exo1* and *Sgs1* on the linear plasmid. Then the proper interpretation would be that the plasmid was too short and its degradation occurred before the BIR mechanism could be completed (Marrero VA *et al.* 2010). More experiments need to be done to confirm that BIR can be inhibited by *Exo1* and/or *Sgs1*. Due to the similarities between BIR and ALT mechanism it is possible that *EXO1* could play a key role in the telomere maintenance by ALT in human cells.

1.8.2 EXO1 in Mismatch Repair

The role of the DNA MMR system is to maintain DNA replication fidelity by repairing DNA mismatches and insertion/deletion loops acquired during DNA replication. And this process is highly conserved from *Escherichia coli* to mammals. Furthermore, the MMR proteins contribute to G1 and G2 checkpoints, the normal apoptotic response initiated by DNA damage, and the fidelity of genetic recombination (Felton KE *et al.* 2007). A lack of MMR contributes to the development of malignancy via several different mechanisms (Felton KE *et al.* 2007). In mammals, the MMR system is formed by the proteins MSH2, MSH3, MSH6, MLH1, MLH3, PMS1 and PMS2. Monoallelic germ line mutations in *MLH1*, *MSH2*, *MSH6* or *PMS2* underlie the inherited Lynch syndrome, which is a cancer predisposition syndrome also known as hereditary non-polyposis colorectal cancer (HNPCC) (Felton KE *et al.* 2007).

The first evidence to suggest that *Exo1* is involved in the MMR system was obtained in *S. cerevisiae* from the observed interaction of *Exo1* and *Msh2*. Now it is known that *Exo1* directly interacts with *Msh2*, *Msh3*, and *Mlh1* and has been suggested to act in MMR by catalyzing 5' to 3' excision and by stabilizing higher order complexes of MMR proteins, (Tran PT *et al.* 2004). Since *EXO1* interacts with members of the MMR pathway, it has been

proposed that *EXO1* mutations could be a cause of HNPCC. Studies in mice have shown that the knockout animals are sterile and have an increase predisposition to lymphomas. Moreover the cells of the *EXO1*^{-/-} mice show defects in MMR and Microsatellite Instability (MIS) (El-Shemerly M *et al.* 2007). The first study that proposed that mutation of *hEXO1* as a possible cause of HNPCC was published by Wu *et al.*, in 2001. In this study the authors screened all the exons of *hEXO1* for mutations in a group of patients with HNPCC that lacked mutations in the MMR genes. They found 30 mutations, but 16 were considered SNPs because they were found in patients and healthy controls. The remaining 14 were unique to HNPCC patients, eight were missense substitutions and one was a splicing variant which resulted in a truncated protein. Figure 1.10 shows the localization of the potential mutations described by Wu (Wu Y *et al.* 2001).

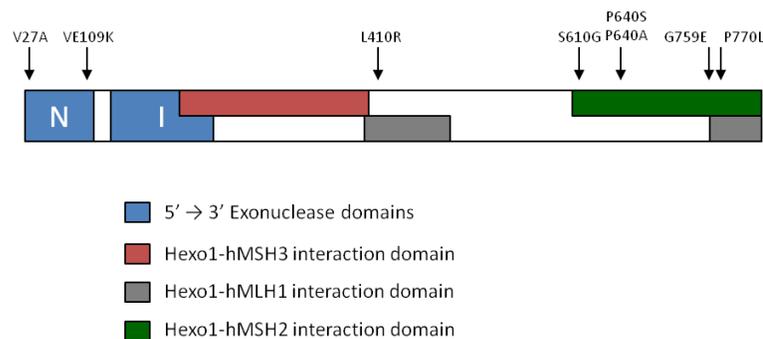


Figure 1.10 Functional domains of EXO1 and localization of the missense mutations identified in patients with HNPCC (Schmutte C *et al.* 2001), (Sun X *et al.* 2002).

However in 2002 Sun *et al.*, evaluated the function of these variants *in vitro* and compared them to the wild type protein. Interestingly the potential mutants retained exonuclease activity, except for E109K and L410R. The E109 mutation is located in the nuclease activity domain, which is highly conserved among eukaryotes cells whereas the L410 mutation is located in the MLH1 interaction domain, and it is also a highly conserved residue among

eukaryotes. The ability of the potential mutants to interact with MLH1 and MSH2 was also tested. Co-immunoprecipitation with MLH1 was similar between the potential mutants and the wild type EXO1. The EXO1 variants P640S, G759E and P770L with base substitutions in the MSH2 interaction domain only precipitated approximately one third of MSH2 protein compared to the wild type EXO1. However, none of the three substitutions completely abolished the interaction with MSH2 (Sun X *et al.* 2002). In 2003 Jagmohan-Cnagur *et al.*, studied another group of HNPCC patients without mutations in the MMR genes and they found in addition to the amino acid variants described by Wu, another 5 potential EXO1 mutations. However, they also found some of the variants from Wu's patients in healthy controls. Interestingly the substitutions that cause loss of enzymatic activity of EXO1 were not found in controls, in support of previous findings. Therefore some EXO1 mutations may be a cause of HNPCC (Jagmohan-Changur S *et al.* 2003). Table 1.3 summarizes the amino acid variants described by both papers in HNPCC patients and the position of each change in the gene and the protein.

In 2005 Yamamoto *et al.*, investigated the role of two SNPs previously described in healthy controls by Wu *et al.*, for their potential role on predisposition, progression and metastasis of colon cancer. They studied the incidence of T439M at exon 10 in the MLH1 interaction domain and the P757L at exon 13 in the MSH2 interaction domain mutations. The study population was patients with colorectal cancer, compared with healthy controls. The results suggested an association between the presence of the M439 and L757 alleles with increased risk of develop colorectal cancer, however the authors stressed that further studies on larger populations are needed in order to confirm such association (Yamamoto H *et al.* 2005). In summary the data that propose base substitutions in the *EXO1* gene as a possible cause of HNPCC is equivocal. At this moment there is not enough evidence to consider mutations of *EXO1* as cause of HNPCC.

In addition to the role of Exo1 in the MMR system, recent data published, in yeast, suggest that Exo1 could participate in the post-replication repair (PRR) pathway in eukaryotes

(Tran PT *et al.* 2007). This pathway facilitates the bypass or tolerance of fork stalling events but does not repair these lesions. In yeast an *Exo1* missense mutant that was functional in PRR but not in MMR was identified. This implies the existence of a domain of *Exo1* required for PRR distinct from those required for MMR (Tran PT *et al.* 2007).

Table 1.3 Variants found in the sequence of *hEXO1*. *Italics* variants just found in HNPCC patients in both studies.

*Variants which show loss of EXO1 enzymatic activity (Sun X *et al.* 2002).

Exon	Nucleotide change	Protein Change	Reference
2	T to C	<i>Val27Ala</i>	(Wu Y <i>et al.</i> 2001)
4	G to A	<i>Glu109Lys*</i>	(Wu Y <i>et al.</i> 2001)
5	G to T	<i>Ala137Ser</i>	(Jagmohan-Changur S <i>et al.</i> 2003)
9	T to G	<i>Leu410Arg*</i>	(Wu Y <i>et al.</i> 2001)
10	T to G	<i>Phe438Cys</i>	(Jagmohan-Changur S <i>et al.</i> 2003)
11	A to G	Ser610Gly	(Wu Y <i>et al.</i> 2001)
11	C to T	Pro640Ser	(Wu Y <i>et al.</i> 2001)
11	C to G	Pro640Ala	(Wu Y <i>et al.</i> 2001)
11	C to T	Asp661Asp	(Jagmohan-Changur S <i>et al.</i> 2003)
13	IVS12-1	G to C	(Wu Y <i>et al.</i> 2001)
13	IVS12-1	G to A	(Jagmohan-Changur S <i>et al.</i> 2003)
13	G to A	Gly759Glu	(Wu Y <i>et al.</i> 2001)
13	C to T	<i>Pro770Leu</i>	(Wu Y <i>et al.</i> 2001)
14	C to T	<i>Ala827Val</i>	(Jagmohan-Changur S <i>et al.</i> 2003)

1.9 CtIP

The human *CtIP* gene is located at 18q11.2 and encodes an 897 amino acid nuclear protein which is significantly conserved among mammals. CtIP is widely expressed in human tissues with testis and thymus showing the highest expression (Wong AK *et al.* 1998). CtIP was first recognized for its role in regulating transcription by interacting with the C-terminal binding protein (CtBP), which is a known repressor of gene expression. CtIP was also identified as a factor able to bind to the pRB family members pRB, p107 and p130. Finally CtIP was also found as an interacting partner of BRCA1. Interestingly some mutations in the *BRCA1* gene which are associated with cancer are able to disrupt the interaction between BRCA1 and CtIP. The discovery that CtIP interacts with proteins involved in tumor suppression raised the idea that *CtIP* could be closely related to cancer development, maybe through a tumor suppressor mechanism (Fusco C *et al.* 1998, Schaeper U *et al.* 1998, Wong AK *et al.* 1988). However one study carried out a survey of 89 tumor derived human cell lines from different origins and they did not find evidence for loss or mutation of the *CtIP* gene (Chinnadurai G 2006).

Recent studies have demonstrated that mutations in the *CtIP* gene can cause one of the forms of Seckel syndrome (SS) linked to 18q11.31-11.2. The patients in one family with Seckel Syndrome showed a homozygous mutation that resulted in a C-terminal truncated protein (Qvist P *et al.* 2011). In addition, a CtIP mutation was identified in another family diagnosed with a less severe form of SS, called Jawad Syndrome. The mutation comprised a two base pair deletion which generated a change in the open reading frame and resulted in alternative splicing of the transcript. Interestingly studies in cells from these patients revealed the presence of both, the aberrant and normal form of the protein, suggesting that the abnormal protein has a dominant-negative effect. This is an unusual mechanism for a recessive disease. Compared to the normal population, SS patients are at higher risk of tumor development, but not as high as for patients with BS or WS. More studies are being carried out to investigate the mechanism that leads to the clinical

manifestations after CtIP mutations (Qvist P *et al.* 2011).

1.9.1 CtIP and DSB repair

The first insight about a possible role of CtIP in DDR came from the observation of ATM-dependent phosphorylation of CtIP after ionizing radiation (IR)-induced damage (Li S *et al.* 2000). In 2007 Sartori *et al.*, investigated the participation of CtIP in DSB repair and observed that after laser microirradiation of cultured human cells CtIP was recruited to the damage sites just after foci of γ H2AX were formed. This observation was found in cells positive for cyclin A, indicating that the response occurs predominantly in S and G2. It was noticed that CtIP co-localized with ATR foci in response to damage. Interestingly, reduced CtIP expression by siRNA, significantly reduced the number of foci of ATR, but not the signal of γ H2AX, suggesting that CtIP is necessary for ATR recruitment after DNA damage. After a DSB is recognized ssDNA is generated and is coated by RPA. Immunofluorescence in cells treated with siRNA against *CtIP* showed reduced accumulation of RPA and ssDNA signal. These effects were reversed after transfection of the cells with a siRNA-resistant plasmid expressing wild type CtIP. This confirmed that these observations were secondary to CtIP depletion. Under normal circumstances, the MRN complex promotes formation of ssDNA. To establish a relation between MRN and CtIP, coimmunoprecipitation experiments were conducted by Sartori AA *et al.* in 2007. These showed co-immunoprecipitation of the MRN complex with CtIP. *In vitro* analysis confirmed that the physical interaction between CtIP and the MRN complex was independent of DNA or other proteins. To investigate the functional relationship of CtIP and the MRN complex, nuclease assays were carried out. It is known that MRE11-RAD50 complex possesses endonuclease activity in the presence of manganese but not magnesium. Using the ssDNA circular genome of Φ X74 it was demonstrated that CtIP not only enhanced the endonuclease activity of MRE11-RAD50 in the presence of manganese but also promoted nuclease activity in the presence of magnesium. Interestingly, CtIP on its own did not

show any detectable nuclease activity. By using a plasmid carrying a reporter that allows evaluation of the frequency of HR events it was observed that depletion of MRN or CtIP reduced the number of HR events. The effect was the same when genes were downregulated separately or at the same time, suggesting that the MRN complex and CtIP cooperate in the HR pathway through a common mechanism. These experiments prompted a search for proteins with similar function in other eukaryotes. The C-terminal 108 amino acids of CtIP were found to be conserved in just one protein in *S. cerevisiae*, the Sae2 enzyme. Sae2 is known to interact with Mre11. Homology analysis of this C-terminal domain led to the conclusion that *CtIP* is the homolog of *Sae2* (Sartori AA *et al.* 2007). However there is an important difference to bear in mind. Functional analysis of Sae2 has shown that this enzyme exhibits endonuclease activity, and this activity can be exerted with or without the presence of the MRX complex (Lengsfeld BM *et al.* 2007). On the contrary CtIP does not have any identified endonuclease activity.

Another important role of CtIP in the HR repair pathway is the recruitment of EXO1. In cells treated with micro-irradiation the recruitment of EXO1 to the sites of DSB was seriously impaired after downregulation of CtIP. A similar effect was observed upon depletion of MRE11. Furthermore, CtIP was detected in immunoprecipitates obtained with anti-EXO1 antibodies. *In vitro* analysis with recombinant proteins confirmed a direct interaction between them (Eid W *et al.* 2010). *In vitro* experiments have shown that CtIP can regulate EXO1 activity: when CtIP was added to the substrate before EXO1, it stimulated the enzymatic activity of EXO1. But when it was added to the substrate after EXO1 the activity of EXO1 was reduced (Eid W *et al.* 2010). The evidence suggests that CtIP play a central role in the HR repair pathway and therefore might be involved in the ALT pathway. Findings about CtIP and its possible role in the ALT pathway will be described in chapter 5.

1.10 Human herpesvirus 6

Telomeres seem to be subject to HR not only to maintain their length and stability, but also HR is likely to be a mechanism used by some viruses to integrate their genomes into human chromosomes. This seems to be the case for the Human herpesvirus 6 (HHV-6) which was first isolated, from peripheral blood mononuclear cells. In 1986 Salahuddin *et al.*, isolated HHV-6 from patients with lymphoproliferative disorders, some of whom were also infected with (Human immunodeficiency virus) HIV or (Human T-cell lymphotropic virus) HTLV (Salahuddin SZ *et al.* 1986). The viral particles showed structural similarities to human herpes viruses but were incapable of infecting T or B lymphoblastoid cell lines. Serological tests demonstrated no immunological cross reactivity with known herpes viruses and these findings lead to the conclusion that this was a novel human herpes-like virus. In 1987 Ablashi *et al.*, classified the new virus as human herpesvirus 6 (HHV-6) (Ablashi DV *et al.* 1987). More strains were isolated from different patients in laboratories around the world and based on their biological and molecular characteristics, two groups of HHV-6 were distinguished: HHV-6A and HHV-6B (Aubin JT *et al.* 1991) (Ablashi D *et al.* 1993).

1.10.1 Organization of HHV-6 genome

The HHV-6 genome is a linear, double stranded DNA molecule. The genome of HHV-6B is 162 kb long containing 119 unique open reading frames (ORFs), nine of which are absent in the A variant. The HHV-6A genome is 159 kb long and contains 119 unique ORFs nine of which have no counterpart in the B subtype (Gompels UA *et al.* 1995). The genomic organization of both variants comprises two major regions: the unique long region (U_L) which spans ~143 kb (Figure 1.11). The genes located in this region are named U1 to U100 and are organized in seven major core blocks conserved amongst herpesviruses (Arbuckle JH *et al.* 2011). The remaining regions of the HHV-6 genome are the left and right direct repeats (DR_L and DR_R), which flank the U_L region and each one spans ~8 kb (Figure 1.11).

ORFs in the DR are designated DR1 to DR8. At the right end of both repeats there is a region of perfect telomere repeats (TTAGGG) called T2 and at the left end a region of imperfect telomere repeats called T1. The DRs also contain the cleavage-packaging motifs, *pac-1* and *pac-2*. These motifs are upstream and downstream of the T1 and T2 elements respectively and play an important role in the generation and packaging of new viral particles (De Bolle L *et al.* 2005).

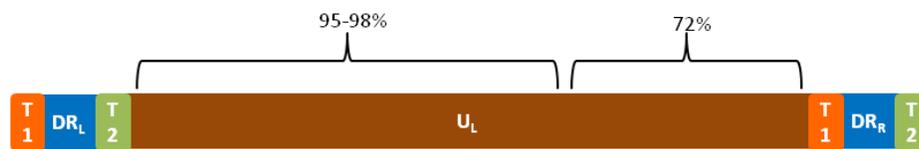


Figure 1.11. HHV-6 genome organization. HHV-6 genome consists of a unique long region (UL) flanked by two direct repeats (DR). Within each DR is a region of perfect telomere repeats (T2) and a region of imperfect telomere repeats (T1). The numbers on top indicate the percentage of sequence identity between the A and B subtype along the UL region. Modified from (De Bolle L *et al.* 2005).

Sequences from both A and B subtypes are 90% identical overall at the nucleotide level. In the UL region the nucleotide identity is between 95 to 98% close to the left end of the virus, and it decreases to 72% towards the right end of the virus, with the exception of some conserved ORFs such as U94. The major divergence is in the direct repeats and the region from ORF U86 to U100 (Dominguez G *et al.* 1999). Some of these differences between regions could explain the biological dissimilarities between A and B variants. For example, the highest degree of divergence is found in the IE-1 region. The splicing pattern and expression of IE-1 differ between the subtypes. The product of the U100 gene, which is involved in CD64 receptor binding and subsequent membrane fusion, shows a considerable degree of divergence. This divergence might contribute to the different tropism observed in each variant, since the A subtype is more commonly found in the central nervous system (CNS) than the B subtype (De Bolle L *et al.* 2005). Variation is found not only between subtypes, but also between different isolates from the same

variant. In the case of the B subtype the highest variability is close to the left end of the DRs (De Bolle L *et al.* 2005). Another intra-strain variation that has been observed is the number of telomeric repeats which can vary from 15 up to 180 repeats in the T2 element (Achour A *et al.* 2009).

1.10.2 Clinical diseases associated to HHV-6 infection

Serologic tests have found that around 90% of children are positives for HHV-6 at age of 13 months (Okuno T *et al.* 1989). In 1988 Yamanishi published that HHV-6B primary infection was a cause of exanthema subitum (roseola infantum) in young children, generally in the first 3 years of life. Clinical symptoms of this disease include high fever, diarrhea, and a mild skin rash along the trunk, neck and face. Isolates from patients with exanthema subitum (ES) are mainly B type (Yamanishi K *et al.* 1988). Despite its high incidence, just 17% of the children develop ES and the majority only present an undifferentiated fever with no rash. The most common complications are inflamed tympanic membranes, gastrointestinal and respiratory symptoms (Caselli E *et al.* 2007). Febrile seizures caused by HHV-6 infection represent between 10-20% of febrile seizures in children under 2 years old. More severe complications like meningoencephalitis and encephalopathy have also been reported but they are rare. HHV-6 can be isolated from cerebrospinal fluid (CSF), which means that direct infection of the central nervous system CNS can occur (De Bolle L *et al.* 2005). HHV-6 neurotropism has been proven as the virus can be detected in different regions of the brain. This means that the CNS is an important place for active and latent infection. Some studies have found a higher prevalence of the A variant in CSF, suggesting that the A variant may infect the brain more easily (De Bolle L *et al.* 2005). Neurological disorders such as encephalitis, ataxia, seizures and multiple sclerosis, have been associated with HHV-6 infection, but there is not enough evidence to confirm HHV-6 as the causal agent of these diseases (Arbuckle JH *et al.* 2011).

The fact that HHV-6 was first described in patients with lymphoproliferative diseases prompted the idea of HHV-6 oncogenic potential. Since this observation several reports have found the presence of HHV-6 in hematologic malignancies (De Bolle L *et al.* 2005). The observation that some fragments of HHV-6 DNA could transform established NIH 3T3 cells and human epidermal keratinocytes *in vitro*, and that these cells could generate metastatic tumors when injected into nude mice supported a role for HHV-6 in cancer development (Razzaque A 1990). Some clinical trials have reported the detection of HHV-6 genome in samples from lymphomas, but different authors argue that this could be a reflection of a latent state and not necessarily prove a relationship between the virus and cancer development. To date there is not enough evidence to confirm or dismiss HHV-6 as a risk factor for cancer development. Further studies to detect viral antigens or viral DNA in large cohorts are needed in order to identify a possible etiological role of the virus in human neoplasias (Ogata M 2009).

Some authors consider HHV-6 as an opportunistic pathogen in immunocompromised patients. For example, in AIDS patients HHV-6 has been associated with fatal cases of pneumonitis and encephalitis (Caselli E *et al.* 2007). In organ transplant recipients, HHV-6 is considered an important factor in morbidity, with infection occurring in almost 50% of bone marrow transplants and in 15-20% of solid organ recipients 2-3 weeks after the procedure. Most of the cases are due to reactivation of latent viruses rather than primary infection and symptoms can vary from fever to transplant failure. Despite the high frequency of infection, the pathophysiologic mechanism underlying reactivation of HHV-6 is not very clear and the role of HHV-6 in post-transplant complications is still under debate (Caselli E *et al.* 2007).

Several groups have investigated whether HHV-6 infection can be a cause of pregnancy loss. Buhtori *et al.*, reported two cases (2.7%) in a group of 48 fetal deaths not related to immunological causes. One was associated with *hydrops fetalis*, which is a generalized accumulation of liquid secondary to malformation of the lymphatic system, and Trisomy

21 and the second with small placenta and growth restriction (Al-Buhtori M *et al.* 2011). Interestingly HHV-6 had been previously found associated with cases of hydrops fetalis with aneuploidies (Ashshi AM *et al.* 2000). Whether HHV-6 can affect pregnancy progress is yet to be determined, but the preliminary findings suggest it could be a factor in some cases of pregnancy loss.

To date, there is no available vaccine against HHV-6. As primary infection is mostly in immunocompetent children, any required treatment is focused on immunocompromised patients. Prophylactic treatment against HHV-6 in transplant recipients is the same as that used against Human cytomegalovirus (HCMV). This consists in nucleoside analogs such as ganciclovir, acyclovir, the nucleotide analogue cidofovir, and the phosphate analogue foscarnet (De Bolle L *et al.* 2005).

1.10.3 HHV-6 Replication Cycle

The replication cycle of the herpesviruses depends on the host cells and biological features of each virus, however some process are shared within the group. During lytic infection the viral envelope fuses with the cell membrane and the linear genome of the virus is released from the capsid into the nucleus. The genome of the virus becomes circular and the expression of the immediate early (IE) genes begins. The expression of these genes mediates the activation of the viral DNA polymerase and the viral genome is replicated via a rolling-circle mechanism. This method of replication leads to the formation of concatemers of viral genomes in a head to tail conformation and cleavage of the concatemers produces single copies of viral genomes. At the last stage of the cycle, other components of the virion are produced in order to generate new viral particles, the cycle concludes with the death of the cell through lysis (Figure 1.12) (Arbuckle JH *et al.* 2011). In cultured cells new viral particles of HHV-6 can be found 6-10 days after infection during the lytic cycle (Krueger GR *et al.* 2003).

However, it is known that not all infections progress to the lytic cycle. A well defined characteristic of herpesviruses is their ability to establish latency within the host cell after circularization of the linear genome to form a nuclear covalently closed circular (ccc) episome (Flamand L *et al.* 2010). Viral expression of latency associated transcripts in conjunction with the cellular DNA polymerase promotes replication and persistence of viral genome without the production of new virus particles. Initiation of the lytic cycle after latency is driven by amplification of the virus genome by viral polymerases through rolling circle mechanism and progression of the productive phase (Arbuckle JH *et al.* 2011).

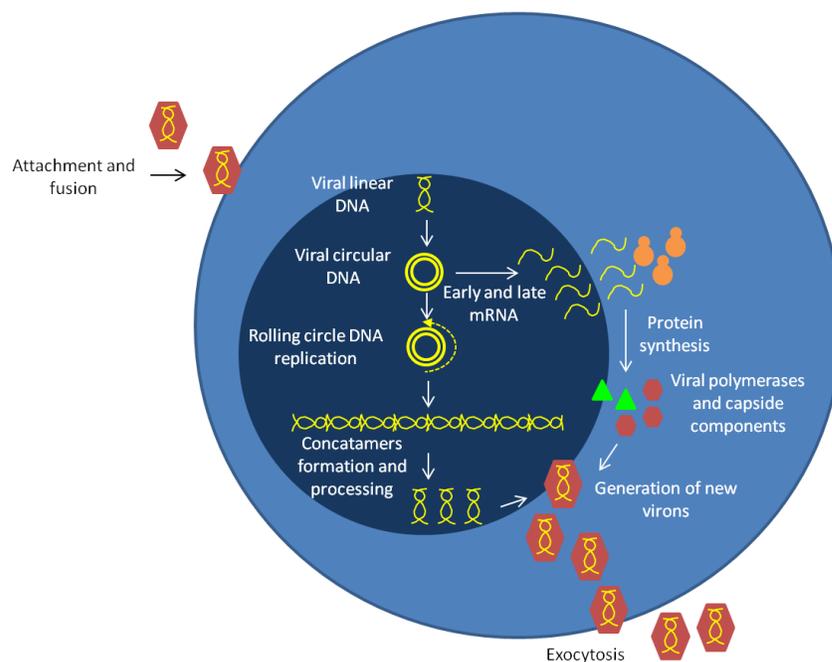


Figure 1.12 Simplified diagram of the lytic cycle of human herpesviruses. The viral envelope fuses with the cell membrane and the linear genome of the virus is released from the capsid into the nucleus. The genome of the virus becomes circular and the expression of the early and late genes begins. The expression of these genes mediates the activation of the viral DNA polymerase and the viral genome is replicated via a rolling-circle mechanism. This method of replication leads to the formation of concatemers of viral genomes. Cleavage of the concatemers produces single copies of viral genomes. At the last stage of the cycle, other components of the virion are produced in order to generate new viral particles.

1.10.4 Latent phase and chromosomal integration of HHV-6

After primary infection of HHV-6, the number of copies in peripheral blood of children increases with a subsequent decrease (De Bolle L *et al.* 2005). However in some patients persistence of high copy number of viral genomes, in the order of 1 million per ml of blood, can be found. In 1993 Luppi *et al.*, reported the presence of unexpectedly high numbers of HHV-6 genomic sequences in circulating PBMCs from patients with no clinical or serological signs of active viral infection. After DNA analysis by Pulsed Field Gel Electrophoresis (PGFE) they found association of the HHV-6 genome with high molecular weight DNA. These observations prompted the idea that the viral genome was associated with cellular DNA. Based on the integration into the human genome of Epstein-Barr virus (EBV) and adeno-associated virus (AAV), these authors proposed that the HHV-6 could integrate into the human genome to achieve latency (Luppi M *et al.* 1993). The first proof of integration of HHV-6 was presented by Torelli *et al.*, in 1995, when they detected by FISH a positive signal in the telomeric region on 17p in three patients with lymphoproliferative diseases (Torelli G *et al.* 1995). After this report, several more papers emerged from different laboratories providing more evidence in favour of the chromosomal integration of HHV-6 at telomeres (Daibata M *et al.* 1998), (Nacheva EP *et al.* 2008). Direct evidence of integration was obtained when Arbuckle *et al.*, were able to amplify by PCR a segment of DNA containing both, human and viral sequences. Specific primers were used for the DR element of the virus and for the human subtelomeric region at 17p, showing that the virus integrated with the DR_L towards the telomere and the DR_R towards the centromere (Figure 1.13) (Arbuckle JH *et al.* 2010). In the same report they also showed the absence of non-integrated episomes in infected cells and they claimed that the virus can be reactivated after treating the cells with the histone deacetylase (HDAC) inhibitor trichostatin A. All the existing data support the idea that a latent state of HHV-6 can be achieved through chromosomal integration. However, some authors think that to accept the integrated HHV-6 as the latent state, it has to be demonstrated that the sequence of the reactivated virus is the same as the integrated one, ruling out the

possibility of a second strain infecting the same individual (Morissette G *et al.* 2010).



Fig 1.13. Orientation of the Chromosomal integrated HHV-6 (CIHHV-6). DR_L and DR_R are orientated towards the telomere and centromere respectively.

The frequency of chromosomally integrated HHV-6 (CIHHV-6) has been studied using qPCR to measure the number of copies of viral genomes per millilitre of whole blood. Ward *et al.*, considered 10^6 copies/ml of whole blood as indicative of CIHHV-6 (Ward KN *et al.* 2006). A survey in the United Kingdom found a CIHHV-6 frequency of 0.8% in healthy donors and a frequency of 2.9% in a group of hospitalised individuals, suggesting a higher frequency in hospital populations. In the same article Leong *et al.*, made a review of previous reports about CIHHV-6 frequency. The frequencies observed ranged from the 0.21% found by Tanaka-Taya *et al.*, in Japanese population (Tanaka-Taya K *et al.* 2004), to 5% reported by Griffiths *et al.*, in 1999 (Griffiths PD *et al.* 1999). Some explanations for this variation could be the number of subjects studied, for example the group from Griffiths included just 25 patients, whereas the group from Tanaka-Taya included 2332. Other considerations are the ethnic group of the subjects studied and the fact that not all suspected positives were confirmed with a different method like FISH (Leong HN *et al.* 2007). More recent articles report a worldwide frequency of CIHHV-6 close to 1%, with a 1:3 ratio of HHV-6A:HHV-6B (Morissette G *et al.* 2010).

After the first indication of the possible CIHHV-6 into human chromosomes, some groups started to look at the possibility of the virus being transmitted through the germline from parents to their offspring. In 1998 Daibata *et al.*, reported a case of a woman diagnosed with acute lymphoblastic leukemia carrying CIHHV-6 at 1q44. After screening her children and grandchildren, they found CIHHV6 at the same locus in one of her two children and one of her two grandchildren. This was the first study claiming transmission of CIHHV-6 in

a Mendelian fashion, with a 50% chance of inheriting the chromosome harbouring the CIHHV-6 (Daibata M *et al.* 1998). One year later Daibata *et al.*, published a case of a woman carrying CIHHV-6 at two different loci, 22q13 and 1q44, inherited from her mother and father respectively (Daibata M *et al.* 1999). With these two reports it was clear that CIHHV6 could be transmitted from one generation to the next. Implying that CIHHV-6 should be present in the germline and that a person can be born with CIHHV-6 in all the nucleated cells of her/his body. These observations have raised many questions about the clinical implications of people being born with 160 kb of virus genome inserted in one telomere.

Following confirmation of the CIHHV-6 questions about the possible mechanisms of integration and whether the integration represents the latent state of the HHV-6 cycle started to arise. Marek's disease affects chickens and is caused by the Marek's disease virus (MDV), which is an oncogenic alpha-herpesvirus that contributes to the development of chicken T-cell lymphomas (Calnek BW 2001). MDV is another example of a herpesvirus that is able to integrate into host telomeres. The MDV genome consists of one unique long (U_L) and one unique short (U_S) region and each U region is flanked by internal and external repeat regions (IRs and TRs respectively). The external repeats (TRs) contain telomeric repeats (TMR). Recently published data have shown that after mutation or deletion of the TMR the virus integrates in multiple intra-chromosomal regions instead of only at telomeres. In addition virus reactivation and tumor development was impaired (Kaufer BB *et al.* 2011). The MDV data in viruses with deleted TMR support the idea of that virus integration into telomeres is mediated by the telomere-like repeats in the viral genome. However, the presence of telomeric repeats might not be enough for virus integration. It is known that human herpes virus 7 (HHV-7) has a similar genome organization to HHV-6, containing T1 and T2 elements (Black JB *et al.* 1999), but to date there is no evidence of chromosomal integration of HHV-7. The arrangement of conserved blocks between HHV-6 and HHV-7 is identical, and the amino acid sequence identity within the core genes is between 41% and 75%, but there is a remarkable difference: neither HHV-7 or any other

herpesviruses possess a homologue of the U94 gene which encodes a homologue of the human adeno-associatedvirus type 2 (AAV-2) Rep78/68 protein (Caselli E *et al.* 2007).

The family of adeno-associated viruses (AAV) represent defective, helper-dependent viruses that need to establish latency to ensure persistence in the host. In humans, latent AAVs persist as episomes, but virus integration has also been detected. The mechanism for integration has been studied in AAV-2 as it has a preferred integration site at 19q13.42. The specificity of integration is mediated by the regulatory protein Rep 78/68 (Huser D *et al.* 2010). Rep 78/68 displays at least three activities related to virus integration: DNA binding, endonuclease activity and helicase-ATPase activity. U94 and rep 78/68 only show 24% amino acid identity, but partial characterization of U94 indicates that it has DNA binding capacity. Any helicase and endonuclease activities have not been identified yet, but key amino acids for these functions are conserved between U94 and rep78/68 (Morissette G *et al.* 2010). In addition Thompson *et al.*, showed that U94 can complement rep-deficient AAV-2, and so proposed that HHV-6 might have acquired a copy of the rep gene from AAV-2 during co-infection *in vivo* (Thomson BJ *et al.* 1994).

Additional information that supports a role for U94 in integration and latency is that U94 transcripts have been detected in PBMCs from HHV-6 positive healthy donors, with no evidence of HHV-6 transcripts associated with the lytic phase of infection. It also has been shown that T cell lines with stable expression of HHV-6 U94 are susceptible to HHV-6 infection, but no signs of lytic infection are seen, suggesting that U94 favours latent phase instead of lytic phase (Rotola A *et al.* 1998). Moreover, U94 is able to inhibit production of HHV-6A and HHV-6B virions in infected T cells in a dose dependent fashion (Caselli E *et al.* 2006). In a similar way U94 can inhibit virion production in fibroblast infected with other herpesviruses such as HHV-7 and HCMV. The mechanism by which U94 inhibit lytic phase of viral infection seems to be related to its DNA binding activity (Caselli E *et al.* 2006). Furthermore, experiments *in vitro* show that the product of the U94 gene binds the human TATA-Binding Protein (hTBP). It has been suggested that this interaction could

facilitate viral control of transcription and could be part of the mechanism to modulate the progress from a latent to a lytic state (Mori Y *et al.* 2000).

In addition to U94 there are other transcripts that seem to be related to the latent stage of HHV-6. Kondo *et al.*, established a system to study HHV-6 latent-infection. This system allowed them to study and compare transcripts from cells with no signs of productive infection with transcripts from cells which develop productive infection. They found that the transcripts from the immediate early 1 (IE1) and immediate early 2 (IE2) regions differed in length depending if they were obtained from cells in latent or lytic phase. In addition, they identified one transcription start site for transcripts from productive cells (productive start site PSS) and at least two transcription start sites for transcripts from cells with no productive infection (latent start site LSS). For each of the LSS they identified one transcript from to the IE1 region and another one from the IE2 region. In total they identified four different transcripts expressed only in the cells with latent infection. These transcripts associated with the latent phase were referred as HHV-6 latency-associated transcripts (H6LTs). Interestingly, they were able to detect the four H6LTs in PBMCs from HHV-6 positive healthy donors; this fact proves that these transcripts are present in naturally infected individuals, not just cells (Kondo K *et al.* 2002). They also studied a group of bone marrow recipients and observed that in the patients with HHV-6 reactivation there was a significant increase in the levels of H6LTs just before the detection of viral activity (Kondo K *et al.* 2003). In the articles published by Kondo *et al.*, they do not make any reference regarding CIHHV-6; new studies using cells with confirmed CIHHV-6 are needed to support the idea of HHV-6 reaching latency through chromosomal integration. So far the mechanisms by which U94 and/or the H6LTs regulate the progression from a latent phase remain to be elucidated.

The current evidence suggests that chromosomal integration could occur by homologous recombination between the T2 element and human telomere repeats. Arbuckle *et al.*, have proposed a model involving the binding of TRF1 and TRF2 to the viral telomere repeats (T2

region) in order to help the virus to become associated with human telomeres. Once the sequences are close enough homologous recombination can take place. U94 could play a role during the integration process, but also in the establishment of latency, something similar to the role of rep78/68 during AVV-2 integration. And they proposed that reactivation of the CIHHV-6 could happen after chromatin remodelling. This idea is supported by the fact that *in vitro* reactivation can be triggered by HDAC inhibitors (Figure 1.14) (Arbuckle JH *et al.* 2011).

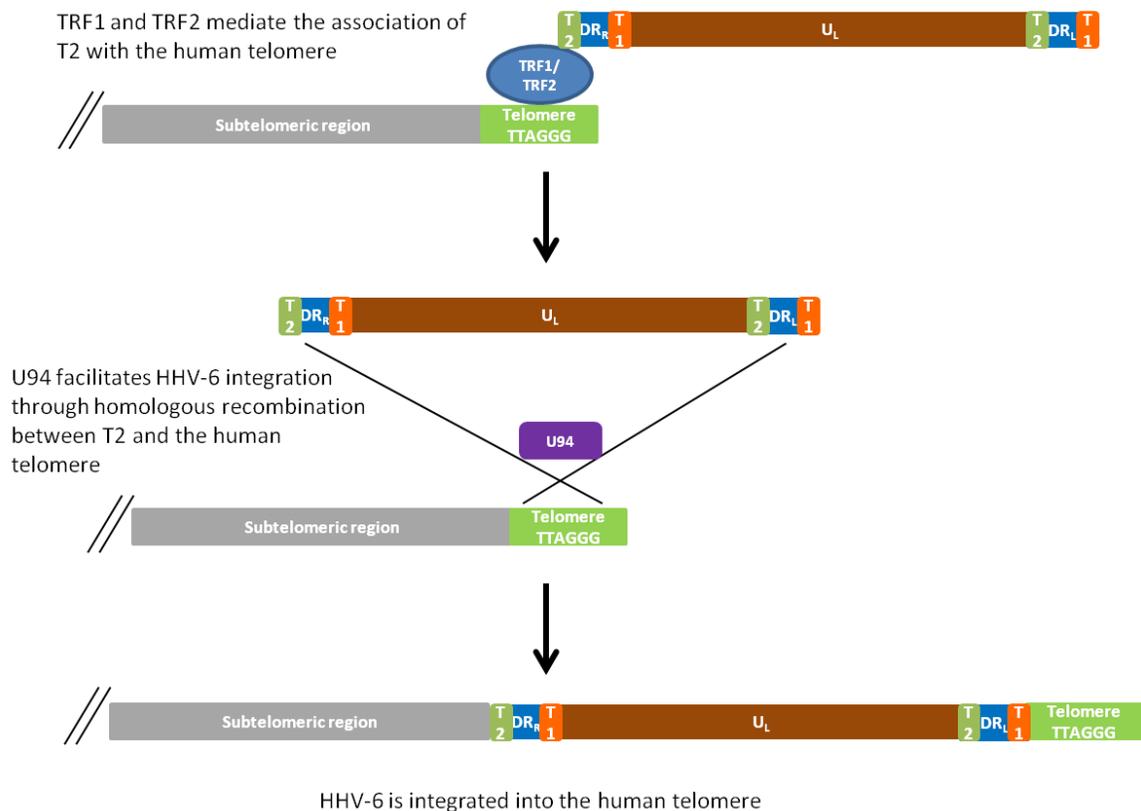


Figure 1.14. Proposed model for CIHHV-6. In the first step TRF1/TRF2 bind to T2 taking HHV-6 close to the human telomere repeats. In the second step U94 promote homologous recombination between viral T2 and human TTAGGG repeats leading to the integration of HHV-6. (Arbuckle JH *et al.* 2011).

1.10.5 Clinical implications of CIHHV-6

Most of the clinical studies regarding HHV-6 do not provide a clear indication whether it is in an integrated form or not. Therefore there is not an accurate picture of the relationship of CIHHV-6 to clinical symptoms or diseases. From now on positive patients for HHV-6 should be carefully screened to look for CIHHV-6, to identify if it is associated with any human pathology. One of the most important areas to investigate is regarding to transplants. In a recent article Lee *et al*, studied the possible influence of CIHHV-6 in the clinical follow up of a group solid organ transplant recipients. They used the criteria of more than one million copies of HHV-6 genome per ml of whole blood as an indicator of CIHHV-6. From the 548 cases studied, seven (1.3%) met this criterion and 35 (6.4%) positives had an average of only 300 HHV-6 copies per ml. They could not identify any symptoms related to active HHV-6 infection. Bacterial infection and transplant rejection were more frequent in the patients with high copy number of HHV-6 genomes but these differences were not statistically significant (Lee SO *et al*. 2011).

Since CIHHV-6 can be found in consecutive generations it is clear that vertical transmission is possible, implying that HHV-6 must reach the germline. It has been published that CIHHV-6 represents the mayor cause of congenital transmission of HHV-6. Among a group of 43 infants with congenital infection 86% resulted from CIHHV-6 (Hall CB *et al*. 2008). This data should warn clinicians about positive cases of congenital HHV-6 and look for CIHHV-6 in relatives.

A recent review published by Pellet *et al*, (Pellett PE *et al*. 2011) outlined issues of clinical relevance regarding HHV-6, they have described different scenarios that could be associated with CIHHV-6, and some of them are: a) Misdiagnosis of active HHV-6 infection in individuals with CIHHV-6. b) Transmission of CIHHV-6 hematopoietic cells from donors to recipients after hematopoietic stem cell transplantation (HSCT). c) Transplantation of a solid organ from and individual carrying CIHHV-6 to a non carrier of CIHHV-6. d) Potential reactivation of CIHHV-6 in immunocompromised host or after treatment with certain

drugs.

In this article they try to offer answers for different scenarios, and they provide valuable information for physicians dealing with patients with CIHHV-6. One relevant point covers virus reactivation. Some authors have published that HHV-6 reactivation is possible *in vitro* by using drugs like the HDAC inhibitor trichostatin A and hydrocortisone (Arbuckle JH *et al.* 2010). However, there is no evidence of reactivation *in vivo*. This is an area that needs to be investigated to avoid the prescription of drugs that can facilitate CIHHV-6 reactivation, especially in immunocompromised patients.

With respect to the immune response it is not known how the presence of the viral genome in every nucleated cell of the body affects the immune system. Tanaka-Taya *et al.*, reported the presence of antibodies against HHV-6 IE protein in 57% of individuals with CIHHV-6 and none in healthy volunteers. However antibodies against HHV-6 Glycoprotein B (gB) were found in 14% of subjects with CIHHV-6 and in 60% of healthy donors. Antibodies against gB play an important role in protections against HHV-6 infection. The explanation could be that because of the presence of HHV-6 in every nucleated cell of the body, and the possible expression of viral proteins during the development of the immune system, these proteins can be recognized as being from the host, leading to immune tolerance (Tanaka-Taya K *et al.* 2004).

Another important topic is regarding the diagnosis of CIHHV-6, Pellet *et al* refer to qPCR as the most reliable and accessible method to detect CIHHV6 and argue that even patients with active infection do not have as high number of copies as those seen in patients with CIHHV-6, however it is possible that some cases of active infections could reach number of copies as high as one million per ml of whole blood, in such cases FISH can help to make the diagnosis. In the future a PCR-based molecular test could unequivocally prove of CIHHV-6, but at the moment at least two problems exist to develop a PCR-based assay. The first is that the virus does not integrate always in the same chromosome, and second is that for most of the chromosomes it is difficult to design specific primers for

subtelomeric regions.

The truth is that HHV-6 shows an unusual behavior for a human herpes virus. Whether it is necessary or not to develop a routine test to screen all newborns to detect CIHHV-6 in order to be aware of risks and/or limitations is yet to be answered. Studies in clinical and basic areas are being carried out to understand the causes and consequences of CIHHV-6 not just for patients, but also for the virus.

1.11 General aim

Human telomeres can be involved in HR-like processes. On one hand HR-like processes help the telomeres to maintain their length in the absence of the enzyme telomerase through the ALT pathway in some types of cancers. On the other hand HR seems to be the way through which HHV-6 integrates its genome into human chromosomes. The main aim of this study is to investigate the mechanisms underlying HR at telomeres in cells that use the ALT pathway and in cells with CI-HHV-6.

1.11.1 Specific aims

The ALT pathway

- To study the role of the RecQ helicases WRN and BLM in the ALT mechanism by analysing the mutation frequency and profile of mutations in a clone from an ALT + cell line that lacks the WRN protein and through the expression of a shRNA, that reduced the level of the BLM helicase.
- To investigate the role of the EXO1 helicase in the ALT mechanism by using shRNA to disrupt expression of this gene in an ALT+ cell line.
- To investigate the role of CtIP in the ALT mechanism by using shRNA to disrupt expression of this gene in an ALT+ cell line.

CI-HHV6

- To determine the frequency of CIHHV-6 in DNA from somatic cells in populations across the world.
- To screen the frequency of CIHHV-6 in DNA from sperm in a panel of British donors.
- To study the effect on telomere length in DNA from individuals with CIHHV-6
- To study the relationship of the CIHHV-6 with the presence of short potentially unstable telomeres in male germ line.
- To analyze the sequence of the Chromosome/HHV-6 junction in order to understand the mechanism responsible of integration.

CHAPTER 2 - MATERIAL AND METHODS

MATERIALS

- Human cell lines

The cell lines that were used in this study are listed in Table 2.6 (see page 72). The WI38 and WI38VA13/2RA cell lines were obtained from European Collection of Cell Cultures (ECACC). The SAOS and U2-OS cell lines were a gift from Dr. P. Salomoni (UCL Cancer Institute; London). The JFCF/6-T.1J/11C, JFCF/6-T.1J/11E and W-V cell lines were kindly provided by Dr. R. Reddel (Cancer Research Unit, Children's Medical Research Institutes, Sydney, Australia). The SUSM1 cell line was kindly given by Professor O. M. Pereira-Smith (Worcester Polytechnic Institute, USA), and the HT1080 cell line was obtained from the American Type Culture Collection (ATCC).

- Human DNA

The panel of human sperm DNA was kindly provided by Professor Alec Jeffreys, HAP/MAP and Human Genome Diversity Project (HGDP) DNA panels were kindly provided by Professor Mark Jobling. The NEOCODEX DNA panel from the Spanish population was kindly provided by Professor Rhona Borts. The DNA from one woman and her two brothers with CIHHV-6 was kindly provided by Professor Martin Dyer.

METHODS

2.1 Description of Single Telomere Length Analysis (STELA)

The Single Telomere Length Analysis (STELA) method (Baird DM *et al.* 2003) is a polymerase Chain Reaction (PCR) based technique that can be used to amplify single telomere molecules in an allele specific fashion. In order to achieve this, a set of three primers is used. An allele specific subtelomeric primer, a primer called Telorette2 which

contains six bases complementary to the telomere 3' overhang with 20 non complementary bases at the 5' end, and the Teltail primer which is identical in sequence to the non-complementary 20bp at the 5' end of Telorette2. The subtelomeric primer contains specific sequence to amplify just one chromosome end and in some instances it is possible to discriminate between alleles of the one pair of chromosomes. The table 2.1 shows the primers sequence used for STELA. The Figure 2.1 shows the principle of this technique.

Table 2.1 Primers used for single telomere amplification (STELA). Sequence, length and annealing temperature of the primers used for STELA.

Primer	Sequence 5'→3'	Length bp	TM °C
Telorette 2	TGCTCCGTGCATCTGGCATCTAACCT	27	Not applicable
Teltail	TGCTCCGTGCATCTGGCATC	20	Not applicable
12qSTELA	CGAAGCAGCATTCTCCTCAG	21	65
XpYpE2	TTGTCTCAGGGTCCTAGTG	19	65
XpYp-415C	GGTTATCGACCAGGTGCTCC	20	65
XpYp-415T	GGTTATCGACCAGGTGCTCT	20	65
17p6	GGCTGAACTATAGCCTCTGC	20	65
HHV6-DR1R	GAAGAAGATGCGGTTGTCTTGT	23	62

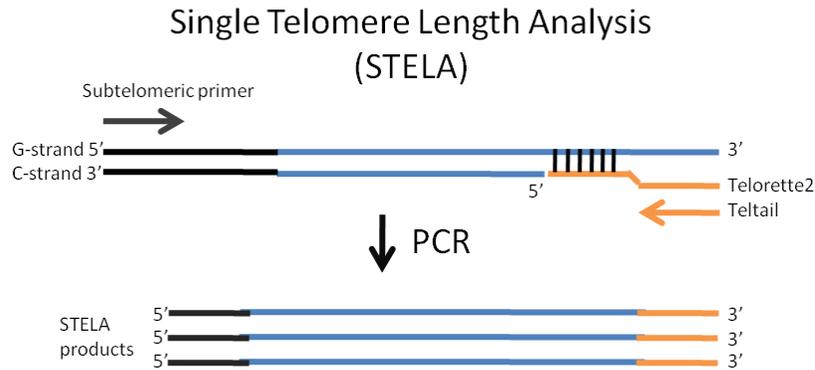


Figure 2.1. The STELA technique. The allele specific subtelomeric primer gives chromosome/allele specificity to the PCR. Telorette2 contains six complementary bases to the 3' overhang, and at its 5' end contains 20 bp that are not complementary to the telomere, but are identical to the sequence of Tetail. In the first round of amplification the telomere is amplified from the subtelomeric primer resulting in copying the sequence of telorette2 at the 3' end of the new synthesised DNA molecule, in this way, Tetail can prime the synthesis of the complementary strand in the subsequent cycles. Each amplicon of different size represent single telomere molecules.

2.1.1 DNA preparation for STELA

From 250 ng to 3 μg of DNA were digested with the restriction enzyme *EcoRI* (NEB) (20U/ μg of genomic DNA) (*EcoRI* does not cut between the subtelomeric primer and the telomere) at 37°C for 5 hrs. The final volume of the digestion was 20-50 μl depending of the initial amount of DNA, after digestion the concentration of DNA was measured in the Nanodrop 1000 (Thermo). The digested DNA was diluted to 10ng/ μl as a working stock. Multiple aliquots (20 μl) at this concentration were prepared to avoid freeze/thaw cycles.

2.1.2 Annealing Telorette2 to the genomic DNA prior to STELA

In 2008 Jeyapalan JN, *et al.* showed that the STELA method is not dependent on the ligation of the primer telorette2 to the C-strand terminus of the telomere as long as the PCRs are conducted without a hot start step. They showed that in the presence of a

telomeric single strand 3' overhang, full length telomere molecules can be amplified by STELA without ligation of the Telorette2 primer to the genomic DNA. Conversely, no STELA products can be obtained in the absence of the telomeric single strand 3' (Jeyapalan JN *et al.* 2008). Therefore all STELA reactions were carried out following annealing (but without ligation) of the telorette2 primer to the genomic DNA in non-hot start PCRs.

Telorette2 was added to the *EcoRI* digested DNA (1ng/ μ l) to allow it to anneal to the telomeric single strand overhangs in a final volume of 10 μ l, the final concentration of the primer Telorette2 was 1mM. Then the DNA/Telorette2 mix was diluted with yeast tRNA (10ng/ μ l) as a carrier, to the desired concentration for STELA. STELA reactions were set up containing 100 to 500 pg of DNA per reaction, 1X PCR buffer (see Appendix 1), 30mM of teltail primer, 300mM of a subtelomeric primer and Taq polymerase (KAPABIOSYSTEMS)/PWO polymerase (Genaxxon) at a ratio of 10:1 (0.1U/ μ l and 0.01U/ μ l respectively) in a final volume of 10 μ l. The PCR cycling conditions included an initial denaturing step at 96°C for 1 min, followed by 25 cycles with denaturation at 96°C for 20 sec, annealing at 65°C for 30 sec and extension at 68°C for 10 min.

2.1.3 Size separation of PCR products using agarose gel electrophoresis

High Gelling Temperature (HGT) agarose (Lonza SeaKem) which provides high gel strength was used for size separation of STELA products. Agarose gels (0.7% HGT) were prepared with 0.5X Tris/Borate/EDTA (TBE) buffer. The percentage of agarose and the size of the gel depended on the desired resolution. To get a good separation of the DNA fragments up to 30kb, long gels (40cm) were used. The electrophoresis voltage and time depended on the size of the gel and resolution desired. For the best separation, 40 cm gels were run at 160V for at least 15hrs in 0.5X TBE containing 0.5 μ g/ml of ethidium bromide. STELA products were separated with two size makers a 1Kb ladder (250bp to 10kb, Fermentas) and High Range Ladder (10.1kb to 48.5kb, Fermentas) on each side so the sizes of the

amplicons could be estimated.

2.1.4 Re-amplification of STELA products

When a single STELA product needed to be re-amplified, a 1:10 dilution of the STELA reaction was made in distilled water (D-H₂O) and one μ l of this dilution was used to set up a re-amplification PCR reaction containing 1X PCR buffer, 200nM of Teltail primer, 200nM of subtelomeric primer and 0.5U of Taq polymerase in a final volume of 10 μ l. Cycling conditions included an initial denaturation step at 96°C for 1min, followed by 30 cycles at 96°C for 20 sec, 65°C for 30 sec, 68°C for 5min. The re-amplified products were size separated in 0.8% HGT agarose gels with 1kb ladder (Fermentas) as a size marker and then purified using Zymoclean™ Gel DNA Recovery Kit (according to manufacturer protocol) and preserved for sequencing (see below).

2.2 Southern Blot preparation and hybridization

2.2.1 Blotting and Hybridization

In preparation for blotting, the agarose gel was washed twice, 7 minutes each, in depurination buffer (0.25M HCl). It was then washed twice, 10 minutes each, in denaturation buffer (0.5M NaOH, 1M NaCl) and afterwards twice, 10 minutes each in neutralizing solution (0.5M Tris, 3M NaCl pH 7.5). The tray for transfer was prepared with 20X sodium chloride/sodium citrate buffer (20X SSC, 3M sodium chloride and 300mM trisodium citrate adjusted to pH 7.0 with HCl). One sheet of 3MM Whatman paper was soaked in 20X SSC and put onto a glass plate over the tray containing 20X SSC, (all the bubbles were removed). A sheet of nylon membrane (Osmonics) cut to the size of the gel and two of 3 MM Whatman paper to fit the gel were pre-wet in 5X SSC. The Southern blot transfer tray was assembled by placing the gel on top of the tray, bubbles were removed

with a glass pipette with a rolling action, and then the nylon membrane was placed on top of the gel, after removing bubbles. Then the two sheets of 3MM paper were placed on top to facilitate capillary transfer of the DNA to the nylon membrane. A five cm pile of paper towels and a glass plate with a 300g weight were placed on top. Paper towels were changed every 15 minutes during the first hour and every 45 minutes during the next four hours. The blot was left to proceed for at least five hours and then it was dismantled; the membrane was baked at 80°C for 10 minutes and afterwards the DNA was crosslinked to the membrane by UV light exposure for 1 minute (short wave length 254nm) in the CL-1000 UV crosslinker (UVP). Two pieces of mesh to fit the membrane were cut and pre-wet in water, the membrane was rinsed in water and put between the mesh pieces. This was then placed in a hybridization bottle with 20ml of Modified Church Buffer (0.5 M Na₂HPO₄, 1 Mm EDTA, 7% Sodium Dodecyl Sulphate) for 30 minutes at 65°C for pre-hybridization to reduce the background signal. Afterwards the buffer was changed and the denatured radiolabelled probes were added and left to hybridise overnight at 65°C with constant rotation.

2.2.2 Post-hybridization washes

The membranes were washed with 0.2X SSC, 0.1% SDS (Sodium Dodecyl Sulfate) at 65°C for 20 minutes three times, (or more if needed) until excess radioactivity was removed. The membrane was rinsed in D-H₂O and dried with whatman paper. The membrane was then placed on a sheet of whatman paper and covered with Saran wrap and then exposed to a phosphor imaging screen in a cassette overnight at room temperature. The screen was scanned with the Typhoon Scanner (GE Health Care). Length analysis was done using Image Quant software (Amersham).

2.2.3 DNA probe labelling for Southern blot hybridization

The telomere double stranded DNA probe and the appropriate size ladder probe were labelled separately with $^{32}\text{P}\alpha\text{dCTP}$, by primer extension using random hexamers. Each reaction contained 5X Oligo labelling buffer (OLB) (see Appendix 1) 6 μl , Klenow fragment of DNA polymerase I 7.5U, $^{32}\text{P}\alpha\text{dCTP}$ 1.5 μl (0.555 Mbq), probe DNA (10ng/ μl) 1.5 μl , bovine serum albumina (BSA) 10ng/ μl 1.2 μl , in a final volume of 30 μl . The reactions were incubated overnight at room temperature. To recover the probe, the reaction was transferred to a glass tube, 50 μl of Oligo Stop Solution (OSS 20mM NaCl, 20mM Tris-HCl pH 7.5, 2mM EDTA pH 8, 0.25% SDS, 1 μM dCTP), 30 μl of high molecular weight salmon sperm DNA (HMWOSS) as a carrier, 25 μl of 2M sodium acetate (NaOAc), and 470 μl of 100% ethanol were added. The precipitated DNA was washed with 80% ethanol, and then resuspended in 500 μl of D-H₂O. The probes were heated at 100°C for three to five minutes to denature the labelled probe and then they were added to the Southern hybridization solution.

2.3 Telomere Variant Repeat mapping by PCR (TVR-PCR)

TVR is a PCR-based technique, previously described (Baird DM *et al.* 1995), which determines the order of the telomere and variant repeats in a single telomere from genomic DNA or STELA products. The reaction uses a forward subtelomeric primer specific for one allele that was end labelled with ^{32}P and a reverse primer that anneals specifically to one of the sequence variant repeats. PCR reactions for each variant repeat were prepared separately. The products were size separated on a 6% denaturing polyacrylamide gel, so that the order of the repeats could be read. The figure 2.2 shows the principle of this technique.

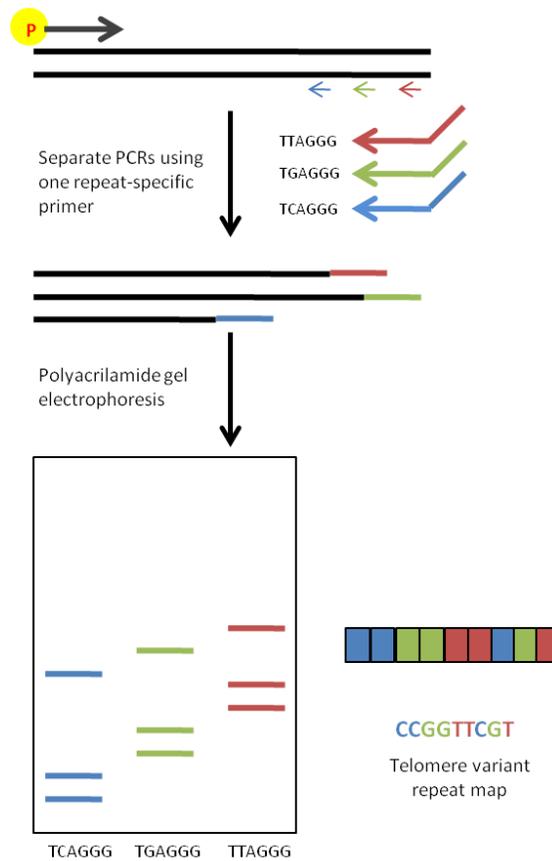


Figure 2.2. TVR-PCR technique. A ³²P end labelled subtelomeric primer is used as forward primer in a PCR reaction with one reverse primer that specifically recognizes only one telomere variant repeat. Amplicons of different size are generated and are size separated on a 6% denaturing polyacrilamide gel using one lane for each variant-specific reverse primer. The array of bands allows the sequence of the repeats in the molecule to be determined.

The table 2.2 lists the telomere variant repeats for which specific primers have been developed and the table 2.3 shows the sequence of the set of primer used for TVR.

Table 2.2 TVR primers. Each primer was designed to recognize specifically one of the telomeric repeats.

Primer	Sequence recognized
TAG-TEL W	TTAGGG
TAG-TEL X	TGAGGG
TAG-TEL Y	TCAGGG
TAG-TEL J	TTGGGG
TEL CTA2	CTAGGG

583 Gel Dryer (BIO-RAD) for 2 hours. When the gel was completely dried it was exposed to a phosphor image screen overnight and scanned with the Typhoon scanner (GE Healthcare).

2.3.3 Polyacrylamide gel preparation

A 50cm wide gel (Sequi-Gen GT Sequencing Cell BIO-RAD) was prepared with 110ml of 6% acrylamide in 1X TBE, 110 μ l of 25% ammonium persulfate (APS) and 110 μ l of Tetramethylethylenediamine (TEMED). The mix was injected into the cast, and left to polymerase overnight at room temperature.

2.4 Human DNA extraction from cell lines

Cell pellets harvested from cell lines were resuspended in 300 μ l of 1X SSC (up to 10^7 cells) and 300 μ l of lysis solution (100 mM Tris HCl pH7.5, 100 mM NaCl, 10 mM EDTA, 1% sarkosyl) and 6 μ l of RNase (10 μ g/ μ l) were added. The lysate was incubated with the RNase for 20 minutes at room temperature. Then 3 μ l (20 μ g/ μ l) of proteinase K were added and the lysate was incubated at 55°C for 5 hours. The samples were transferred to a MaXtract column (Qiagen). And an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) was added. Samples were mixed gently to obtain a good emulsion, followed by centrifugation at 12000 rpm for 10 minutes. MaXtract columns contain a gel that helps to separate the phase containing the phenol and proteins from the soluble phase containing the DNA. The upper phase was transferred to a new tube and the organic phase was discarded. To precipitate the DNA, 0.1 volumes from the original volume of 2M NaOAc and 2.75 volumes of 100% ethanol were added and mixed gently. The DNA pellet was transferred to a new tube and washed with 1ml of 80% ethanol,. Then the ethanol was decanted and the pellet dried at room temperature. The pellet was resuspended in the

required volume of 10mM Tris pH 8.0 and left overnight at 4°C to allow it to dissolve.

2.5 Polymerase Chain Reaction and Sequencing

2.5.1 *EXO1*

The primers used to amplify the promoter region and the coding exons of the *EXO1* gene were described previously (Jagmohan-Changur S *et al.* 2003). The primer sequences, annealing temperatures and amplicons size are described in Table 2.4. Each PCR reaction contained 1X PCR buffer, 500nM of forward and reverse primers, 0.5U of Taq polymerase, 50ng of genomic DNA in a final volume of 10 μ l. Cycling conditions included an initial denaturing step at 96°C for 1 min followed by 30 cycles at 96°C for 1 min, 55°C-60°C (depending on the exon) for 30 sec and 72°C for 30 sec. The PCR products were separated by electrophoresis on a 2% agarose gel at constant 100 Volts for 1 hour with the 1kb ladder (Fermentas) as size marker. The PCR products were purified from the gel using the Zymoclean kit (ZYMO research) according to manufacturer's protocol, and then sequenced with the Big Dye Terminator V3.1 ready mix. Each sequencing reaction contained 1 μ l of the Big Dye 3.1 terminator sequencing mix, the sequencing primer 320nM, 20-30ng of the purified PCR product and 1X Big Dye buffer in a final volume of 10 μ l. Cycling conditions were 25 cycles with denaturation at 96°C for 10 sec, annealing at 50°C for 5 sec and extension at 60°C for 4 min.

After the sequence reactions were completed the samples were cleaned with a hot SDS treatment by adding 10 μ l of D-H₂O and 2 μ l of 2.2% SDS, in order to get a final concentration of 0.2% SDS. The samples were heated at 98°C for 5 min and cooled to 25°C for 10 min in thermocycler, then purified using the Performa[®] DTR Gel Filtration Cartridges (EdgeBio), according to manufacturer's protocol. The clean up treatment helped to remove products below 20bp and unincorporated nucleotides from the Big Dye terminator 3.1 mix which can cause unwanted dye peaks on the chromatogram. The

sequences were sent to the PNAFL service for electrophoresis and the sequence reads were analysed using Autoassembler software.

Table 2.4 Primers used for *EXO1* PCR amplification and sequencing. Obtained from (Jagmohan-Changur S *et al.* 2003)

Exon	Sequence 5'→3'	TM°C	Primer length	Amplicon size bp
Promoter	CAGGACGCAACCCTATGAGT	60	20 bp	661
	GACGACACGTTTCCTTAGGC		20 bp	
ex1	GCTGGCAGTCCAGGTTTAC	61	20 bp	428
	TTAATCTGGATGCAGGTGGTC		21 bp	
ex2	GGACTCCAAGCTTCTCTTCA	61	21 bp	429
	GGAAGTCATCCCTGATTTGG		20 bp	
ex3	CATAGTTTTCTCATCTGGCCTAAAA	65	25 bp	355
	GTGATCCTCTTGCCTCAACC		20 bp	
ex4	CAAGGGCCTGGTGTGTACTT	65	20 bp	347
	TCCACCCTTGATTTTCTTCT		21 bp	
ex5	TACTTCCAGATGCCGTGCT	61	20 bp	418
	TCAAGCAGCAAACCTTCT		20 bp	
ex6	TTGCAGTGGATTAGATAATGACAA	57	24 bp	431
	GGCAAAGTTCTAGTGCCTCA		20 bp	
ex7	GCAGTTAATGTTTCAATCCCTCTT	61	24 bp	318
	TGAAGAAAAATGCACAAATTCTACC		25 bp	
ex8	GGATTTTCGTGACAATAGAAAACA	61	24 bp	314
	ATTCTGCACAGAGCCTGGAG		20 bp	
ex9	GGACCAGCAAAGTAGGAAATG	59	21 bp	475
	GAAAATCTTCGCGACTTTGC		20 bp	
ex10	GGCCAAATCTCTAAGTACAGGTG	59	23 bp	337
	AGACAATCAACATTCTGGAGACA		23 bp	
ex11a	TGAATCTTGACACCCCTTGA	61	20 bp	428
	GCTCGGCGTTCTTGAAAA		18 bp	
ex11b	GCAGCAAATTTACAAGGACCA	59	21 bp	439
	TCGTCAATCACATGTAGGAAAAA		23 bp	
ex12	GGCAAATATCATCCTTTCCA	61	20 bp	305
	AGCATACACCGCTTCTGAAA		20 bp	
ex13	TGAGAGTCCATCTCAAAAAGTAAA	61	24 bp	396
	AGGAAGAGTTGGGAGAAAGG		20 bp	
ex14a	CAGGAGAACCGGATTGTGTT	55	20 bp	409
	TTTAAAATGGCACAAGAATGAGAAA		24 bp	
ex14b	ATGTGGCCGTGTTCAAAGA	65	19 bp	436
	ACAGATGGGAAAGGCAAGTG		20 bp	

2.5.2 HHV6

2.5.2.1 Chromosome-HHV6A junction PCR

PCR to confirm the junction between the chromosome and HHV6A in the KUK cell line was performed using the primer 10qF and a primer specific for the DR8 region of the HHV-6A (see Table 2.5 for primers sequence). The 10qF primer was designed in Dr. Royle's laboratory from the sequence obtained by inverse PCR to isolate the chromosome:HHV-6A junction. PCR reactions were set up containing 1X PCR buffer, 500nM of primer 10qF, 500nM of primer DR8F, 0.5U of Taq polymerase, 20ng of genomic DNA in a final volume of 10 μ l. Cycling conditions included an initial denaturation at 96°C for 1 min, then 30 cycles at 96°C for 15 sec, 64°C for 20 sec, 72°C for 1 min. PCR products were size separated in 1% agarose gels prepared in 0.5X TBE by electrophoresis at constant 100 volts for 1 hour, using the 1 kb ladder (Fermentas) as size marker.

2.5.2.2 PCR for HHV6 Screening

Genomic DNA samples from donors were screened for the presence of the HHV6 U18 gene using the primers U18F and U18R that anneal to the U18 region of HHV6A. This region was selected because it is present in both A and B variants and this pair of primers can amplify the region from both. PCR reactions were set up containing 1X PCR buffer, 500nM of primer U18F, 500nM of primer U18R, 0.5U of Taq polymerase, 5ng of genomic DNA in a final volume of 10 μ l. Cycling conditions included an initial denaturation at 96°C for 1 min, then 30 cycles at 96°C for 30 sec, 65°C for 20 sec, 72°C for 1 min. PCR products were size separated on 1.5% agarose gels prepared in 0.5X TBE by electrophoresis at constant 100 volts for 1 hour, using the 1 kb ladder (Fermentas) as size marker.

Once the positives samples for HHV6 were identified, amplification of specific regions of each HHV6 type was carried out in order to determine whether the sample contained the

HHV6-A or B variant. Primers DR5F-DR5R and primers DR6F-DR6R are specific HHV6A and HHV6B respectively (for primer sequences see Table 2.5). PCR reactions were set up containing 1X PCR buffer, 500nM of primer DR5F or DR6F, 500nM of primer DR5R or DR6R, 0.5U of Taq polymerase, 5ng of genomic DNA in a final volume of 10 μ l. Cycling conditions included an initial denaturation at 96°C for 1 min, then 30 cycles at 96°C for 30 sec, 60°C for 20 sec, 72°C for 1 min. PCR products were size separated in 1.5% agarose gels prepared in 0.5X TBE by electrophoresis at constant 100 volts for 1 hour, using 1 kb ladder as size marker.

Table 2.5 Primers used for HHV-6 PCRs. Sequence, annealing temperature, length and amplicon size of the pairs of primers used to amplify the U18, DR5, DR6 and DR8 regions of HHV-6. The 10qF primer was used in conjunction with the DR8F primer to confirm CIHHV-6A in the cell line KUK.

Primer	Sequence 5'→3'	TM°C	Length bp	Amplicon size bp
U18F	CATATCTGATCAACCTTGCGATG	65	23	632
U18R	ATAACAGCATCGTAAATGCACCC		23	
DR5F	CGTCGACTTCTCGTTCTTTATGC	60	23	342
DR5R	CACATATCCATGAACGGACACAC		23	
DR6F	AGGCGTGATTCTGGGAAAC	60	19	200
DR6R	CCGAATACGTCCAATGTCCT		20	
DR8F	GCAGAGACAAAAGTATGCGGAAG	60	23	177
DR8R	GGATTACGGAGGTGAATGTTGC		22	
10qF	ATCCTTCCTCTTTGCAGCCG	64	20	N/A

2.6 Cell lines and cell culture

For each cell line a vial of cells stored in liquid nitrogen was thawed and cultured in 25 cm² flasks with 8ml of media and incubated at 37°C, with 5% CO₂. Table 2.1 shows the media used for each cell line.

Table 2.6 Cell lines used and medium used for each cell line.

Cell line	Cell Type	Telomere-maintenance mechanism (TMM)	Transformation mechanism	Media
W-V	Fibroblast	ALT+	SV-40 immortalized	DMEM + 10% FCS
SAOS	Fibroblast (osteosarcoma), Caucasian female	ALT+	Tumor-derived	RPMI + 10% FCS
WI38VA13/2RA	Fibroblast (lung), Caucasian female	ALT+	SV-40 immortalized	MEM + 1% NEAA + 10% FCS
U-2OS	Fibroblast (osteosarcoma), Caucasian female	ALT+	Tumor-derived	DMEM + 10% FCS
SUSM1	Fibroblast (liver)	ALT+	Chemical	MEM + 1% NEAA + 10% FCS
JFCF/6-T.1J/11E	Fibroblast	ALT+	SV-40 immortalized	DMEM + 10% FCS
HT1080	Fibroblast (Fibrosarcoma), Caucasian male	Telomerase +	EBV	DMEM + 10% FCS
JFCF/6-T.1J/11C	Fibroblast	Telomerase +	SV-40 immortalized	DMEM + 10% FCS
WI38	Fibroblast (lung) female	Normal (Telomerase-/ALT -)	SV-40 immortalized	MEM + 1% NEAA + 10% FCS

When the cells were confluent they were harvested and transferred to an 80 cm² flask with 18ml of media until confluence was reached again. To harvest the cells the media was removed by aspiration, the cells were washed with 5ml of 1X Phosphate buffered saline (PBS), and then incubated in 2ml of 1X Trypsin (Gibco) for 5 minutes at 37°C. After the incubation in trypsin solution, cells were detached by agitation. The suspension of cells was transferred to a 15ml tube with 8ml of fresh media, mixed and centrifuged at 1100rpm for 8 minutes. The media was removed and the pellet washed with 10ml of 1X PBS and centrifuged 8 minutes at 1100 rpm. The PBS was removed and the pellets frozen in dry ice for storage at -80°C.

2.7 Protein Extraction from cell pellet

Lysis buffer (Tris pH 7.6 50 mM, NaCl 150 mM, EDTA 1 mM, Triton 0.5%-1%, SDS 1%), was kept at 4°C. On the day of extraction 10 µl of protease inhibitor cocktail (SIGMA) and 10µl of phenylmethanesulfonylfluoride (PMSF) 100X were added per one ml of lysis buffer. Then 50µl of the Lysis buffer mix were added to the cell pellet (~10⁶ cells). The cell lysate was incubated on ice for 10 minutes and mixed thoroughly. Afterwards the cell lysate was incubated for an additional 10 min and then centrifuged for 20 minutes at 13000 rpm at 4°C to pellet the debris. The supernatant containing the proteins was transferred to a new tube and kept it on ice. The protein concentration was measured using the Bradford method (Bradford MM 1976).

The Bradford method relies on the binding of Coomassie Brilliant Blue (see Appendix 2) to proteins, the method is based on the fact that the maximum absorbance of the Coomassie Brilliant Blue acidic solution shifts from 465nm to 595nm in the presence of proteins. Using the Eppendorf biophotometer (Eppendorf) a standard curve measuring the absorbance at 595 nm (OD₅₉₅) of a series of standard samples with known concentration of BSA was set up. The standard curve was prepared by making mixes of 1ml of the Coomassie Brilliant Blue reagent with BSA at different concentrations ranging from 0 to

20µg/µl. The samples were incubated for 5 minutes at room temperature and then the OD₅₉₅ for each standard was measured using the biophotometer. Dilutions of the unknown samples were then prepared in 1ml of Coomassie Brilliant Blue, mixed and incubated for 5 minutes at room temperature and the OD₅₉₅ was measured. The concentration of the extracted protein of each sample was then determined by comparison with the readings of the standard curve by the biophotometer.

2.8 Protein separation by SDS-PAGE and Western Blot detection

Gels used to separate proteins consisted of a separating gel and the stacking gel on top. A 7.5% polyacrylamide gel provides good separation for proteins up to 100 KDa. An eight cm separating polyacrylamide gel at 7.5% was prepared with protogel (40%, 19:1 acrylamide:bisacrylamide, Geneflow) 1.5ml, lower buffer (1.5M Tris-HCl, 4% SDS pH 8.8) 1.5ml, D-H₂O 3ml, 10% Ammonium persulfate (APS) 75µl, Tetramethylethylenediamine (TEMED) 5µl. The mix was poured into the gel mould and 200µl of isopropanol added to remove air bubbles on the upper edge of the gel. The gel was left to polymerise for 30 min, then, the isopropanol was decanted by inverting the mould and then the mould was rinsed briefly with D-H₂O. The stacking gel (3%) was prepared using 325µl of protogel, 625 µl of upper buffer (0.5 M Tris-HCl, 4% SDS pH 6.8), D-H₂O 1.5ml, 10% APS 75 µl, TEMED 5 µl, mixed and poured into the gel mould on the top of the separating gel. The comb was placed carefully on top and the gel was left to polymerise for 60 min.

The comb was removed, the gel placed in the electrophoresis chamber and the chamber was filled with running buffer (250Mm Tris, 1.92M glycine, 1% SDS). In order to disrupt secondary structures so the proteins would migrate depending on their molecular weight rather than shape, the appropriate amount of 4X Laemmli loading buffer (4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromphenol blue, 0.125 M Tris HCl) was added to a 1X final concentration. The samples were heated for 5 minutes at 65°C. After heating, the samples were loaded into the gel and the wells were completely filled with 1X running

buffer. The centre of the chamber was filled with 1X running buffer and electrophoresis was carried out at a constant 150 Volts for 1 hour.

Six pieces of 3 MM Whatman paper and 1 of nitrocellulose membrane (HyBOND ECL Amersham) that fit the gel were cut. Three pieces of Whatman were soaked in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol) and placed in the transfer apparatus, (TE 70 Semi-dry transfer unit, Amersham). Air bubbles were removed with a rolling action with a glass pipette and the membrane was pre-wet in transfer buffer for 10 minutes and placed on top of the Whatman paper. Then the gel was soaked in transfer buffer and placed on the membrane. The remaining pieces of Whatman paper were soaked in transfer buffer and put on top of the gel. The transfer apparatus was assembled and 1kg of weight was put on top of it. Transfer was carried out at a constant 60mAmp for one hour. At the end of the transfer the Western blot was ready for protein detection.

2.8.1 Protein detection using an antibody and staining

The Western blot membrane was blocked with 5% Marvel in 1X TBS (50mM Tris/HCl pH 7.4, 150mM NaCl/0.3% Tween 20) overnight at 4°C. The membrane was incubated with 4ml of the primary antibody diluted as appropriate in 5% Marvel in TBS/0.3% Tween 20 for 1 hour with shaking. The table 2.7 shows the dilutions for the primary and secondary antibodies.

Table 2.7. Antibodies used for Western blot analysis

Antibody	Type	Dilution	Supplier
IgG goat anti-Exo1	Primary	1:400	Santa Cruz
IgG mouse anti-GAPDH	Primary	1:10000	Santa Cruz
IgG Rabbit anti-CtIP	Primary	1:3000	Abcam
IgG mouse anti-Tubulin	Primary	1:7000	Abcam
IgG donkey anti-goat	Secondary	1:10000	GE
IgG rabbit anti-mouse	Secondary	1:20000	GE
IgG donkey anti-rabbit	Secondary	1:10000	GE

Following the incubation with the primary antibody the Western blot was washed 4 times with 1X TBS/0.3% Tween 20, 5 minutes each. Horse Radish Peroxidase (HRP) conjugated antibodies were used for detection. Five ml of the secondary antibody were added at the appropriate dilution in 5% Marvel, 1X TBS/0.3% Tween 20 and the membrane was incubated for 45 minutes with shaking at room temperature. The western blot was then washed 4 times with 1X TBS/0.3% Tween 20, 15 minutes each. HRP was developed with ECL kit from Amersham Life Science (Cat. RPN2108). According to manufacturer's protocol 100µl of solution B (peroxide solution) were added to 4ml of solution A (luminol solution), mixed and poured over the membrane. The membrane was incubated for 5 minutes in the dark. Excess reagent was removed and the membrane wrapped in Saran and exposed to X-ray film at room temperature. The exposure ranged from 10 seconds to 1 hour depending on the intensity of the signal.

2.9 RNA Extraction

RNA was extracted from frozen pellets prepared from cultured cells. The pellet was resuspended in 1ml of TRI Reagent (Sigma-Aldrich). 1ml of reagent is enough for $5-10 \times 10^6$ cells. The sample was incubated in TRI reagent at room temperature for 5 minutes to allow complete dissociation of the nucleoprotein complexes. Chloroform was added (0.2ml per 1ml of TRI Reagent used), shaken vigorously for 15 seconds, followed by an incubation for 10-15 minutes at room temperature. The sample was centrifuged at 12000 x g for 15 minutes at 4°C. After centrifugation three phases can be observed: the lower pink phase contains the proteins dissolved in phenol, the middle cloudy phase contains most of the DNA and the clear upper phase contains the RNA. The upper phase was carefully transferred to a new clean 1.5ml tube, avoiding the middle cloudy phase. To precipitate the RNA isopropanol was added (0.5ml per 1ml of TRI reagent used), the tube was shaken, incubated for 5-10 minutes at room temperature and then centrifuged at 12000 x g for 10 minutes at 4°C. The RNA formed a pellet on the wall of the tube. The

supernatant was removed and the pellet washed with 1 ml of 75% ethanol per 1 ml of TRI Reagent used to remove excess salt from the TRI reagent. The sample was vortexed and centrifuged at 7500 x g for 5 minutes at 4°C. The RNA pellet was dried at room temperature for 10-15 minutes, (to evaporate the ethanol, more time can cause RNA degradation). The required volume of RNase free D-H₂O was added and the RNA solution was heated at 55°C for 10 minutes to facilitate dissolution. The concentration of the RNA, the 260:280 and the 260:230 ratios were measured with the Nanodrop 1000 instrument (Thermo Scientific).

2.10 cDNA Synthesis

cDNA was synthesised using the Verso™ cDNA Kit (Thermo) according to the manufacturer's protocol. Each reaction contained 4µl of 5X cDNA synthesis buffer, 2µl of dNTP mix, 1µl of RNA primer (Random primers), 1µl of RT Enhancer (to remove contaminating DNA), 1µl of Verso Enzyme Mix (Verso™ reverse transcriptase and RNase inhibitor), 1µg template (RNA) and D-H₂O up to 20 µl. For each reaction, 1 µg of the RNA template in 11µl of D-H₂O was heated for 5 minutes at 72°C to remove secondary structures. Afterwards 9µl of a mix containing the appropriate volumes of the rest of the reagents were added. The cDNA synthesis reaction was incubated at 42°C for 30 minutes and then the enzyme was inactivated at 95°C for 2 minutes. At the end of the reaction 20µl of D-H₂O were added to each sample in order to achieve cDNA concentration of 25ng/µl.

2.11 Real Time PCR

Relative quantification by Real Time PCR (qPCR) was used to measure the mRNA level of a gene of interest and analyse the effect of the shRNA on the expression of the target genes in the transfected clones. Precision S-Y qPCR Master Mix and primers were acquired from

Primer Design LTD. Primers specific for each gene spanning exon-intron boundaries were designed to work at 60°C and qPCR was carried out in a Lightcycler 480 (Roche). The total amount of cDNA per reaction was 10 ng unless otherwise specified. Data were analysed using qbase^{plus} software from Biogazelle LTD. Primers sequences are shown in Appendix 1.

According to the manufacturer's recommendation each qPCR reaction contained 1µl of reconstituted forward and reverse primers (300nM each), 10µl of 2X Precision S-Y qPCR Master Mix containing SYBR Green and 5µl of cDNA (10ng) in a final volume of 20µl. Cycling conditions consisted of a single initial enzyme activation step at 95°C for 10 min and 40 cycles of denaturation at 95°C for 15 sec followed by annealing/extension at 60°C for 1min.

2.11.1 Reference Gene Selection

A critical step in performing qPCR is the selection of one or more reference genes to normalize the levels of expression of the gene of interest. Vandesompele *et al*, reviewed the necessity for more than one control gene in qPCR experiments, as well as the selection of the most appropriate reference for each sample (Vandesompele J *et al*. 2002). For the selection of the reference genes the GeNorm reference genes selection kit from Primer design was used. The kit includes primers for 12 constitutively expressed genes. The Table 2.8 shows the genes included in the GeNorm reference genes selection kit from Primer design and table 2.9 shows the sequence of the primers used for *EXO1* and *CtIP* as target genes.

Following the primer design recommendations, 8 different RNA samples from each cell line were used to determine which gene showed the most stable expression in the cell line, duplicates of each reaction containing 10ng of total cDNA were set up and the data were analyzed using the qbase^{plus} (Biogazelle) software.

Table 2.8 Genes included in the GeNorm reference genes selection kit from Primer design. Primers for reference genes were obtained from Primer design. The Primers sequence were not provided by the manufacturer.

Homo sapiens actin, beta (ACTB), mRNA.
Homo sapiens glyceraldehyde-3-phosphate dehydrogenase (GAPDH), mRNA.
Homo sapiens ubiquitin C (UBC), mRNA.
Homo sapiens beta-2-microglobulin (B2M), mRNA.
Homo sapiens phospholipase A2 (YWHAZ), mRNA.
Homo sapiens splicing factor 3a, subunit 1(SF3A1), mRNA.
Homo sapiens 18S rRNA gene
Homo sapiens cytochrome c-1 (CYC1), mRNA.
Homo sapiens eukaryotic translation initiation factor 4A, isoform 2 (EIF4A2), mRNA.
Homo sapiens succinate dehydrogenase complex, (SDHA), mRNA.
Homo sapiens topoisomerase (DNA) I (TOP1), mRNA.
Homo sapiens ATP synthase, (ATP5B), mRNA

Table 2.9 Primers used for *EXO1* and *CtIP* real time PCR. Sequences, annealing temperatures, lengths and amplicons size of the primers used for real time PCR for *EXO1* and *CtIP*. Primers were obtained from Primer Design.

Primer	Sequence 5'→3'	TM°C	Length	Prod. size bp
Exo1 Fw	AGTCAGTATTCTCTTTTCATTTACGA	60	25	121
Exo1 Rv	CTCTTTTTGTAAAGTAGGTCCATTTTC	60	26	
CtIP Fw	GATTCACCGATAACAGCCTTCT	60	22	108
CtIP Rv	CAGAGTGCTCCAATTTAGTATGTG	60	24	

2.11.2 qPCR efficiency

The amplification efficiency for each amplicon was assessed from a standard curve using 5 fold dilutions from 100ng to 0.16ng of total cDNA. Efficiencies were calculated using *qbase^{plus}*, values between 1.90 and 2.00 being considered good.

To validate the use of a reference gene, the test and reference genes should show similar amplification efficiencies. In order to determine whether this was true, ΔC_T values were calculated by subtracting the C_T value of the reference gene to the C_T value of the test gene obtained at each dilution point of the standard curve. The ΔC_T values were plotted against the Log_{10} of the cDNA input at each dilution point and the slope of the curve was calculated using Excel. A slope <0.1 was considered as showing similar efficiencies among amplicons (control and experimental amplicons).

2.12 Inverse PCR (IPCR)

Inverse PCR (IPCR) was used to obtain the sequence of the HHV6A-Chromosome junction. Primers IPCR-1 and IPCR-2 published by (Arbuckle JH *et al.* 2010) were designed to prime within the DR8 region of the HHV6 genome, just upstream of the T2 (perfect telomeric repeats) region. The primer sequences for IPCR are shown in Table 2.10. IPCR was described previously by Ochman *et al.*, (Ochman H *et al.* 1988). This technique facilitates the amplification of an unknown DNA sequence that is located next to a known DNA sequence by designing primers within the known sequence orientated in opposite directions. Figure 2.3 shows the principle of this technique.

Table 2.10 Primers used for inverse (PCR IPCR). Sequence, length and annealing temperature of the primers used for IPCR.

Primer	Sequence 5'→3'	Length bp	TM °C
IPCR-1 (HHV6-A)	GCACAACCCACCCATGTGGTAGTCGCGG	28	64
IPCR-2 (HHV6-A)	CGTGTGTACGCGTCCGTGGTAGAAACGCG	29	
IPCR-1seq (HHV6-A)	GGGTACGTAGATGGGGCATA	20	64
IPCR-2seq (HHV6-A)	CAGTCCCGATCCTTCCTTT	20	
IPCRB-1 (HHV6-B)	GAGCATACGGGGGGCAGAT	19	64
IPCRB-2 (HHV6-B)	CTGCCATAGAGGTGGTGTGT	20	

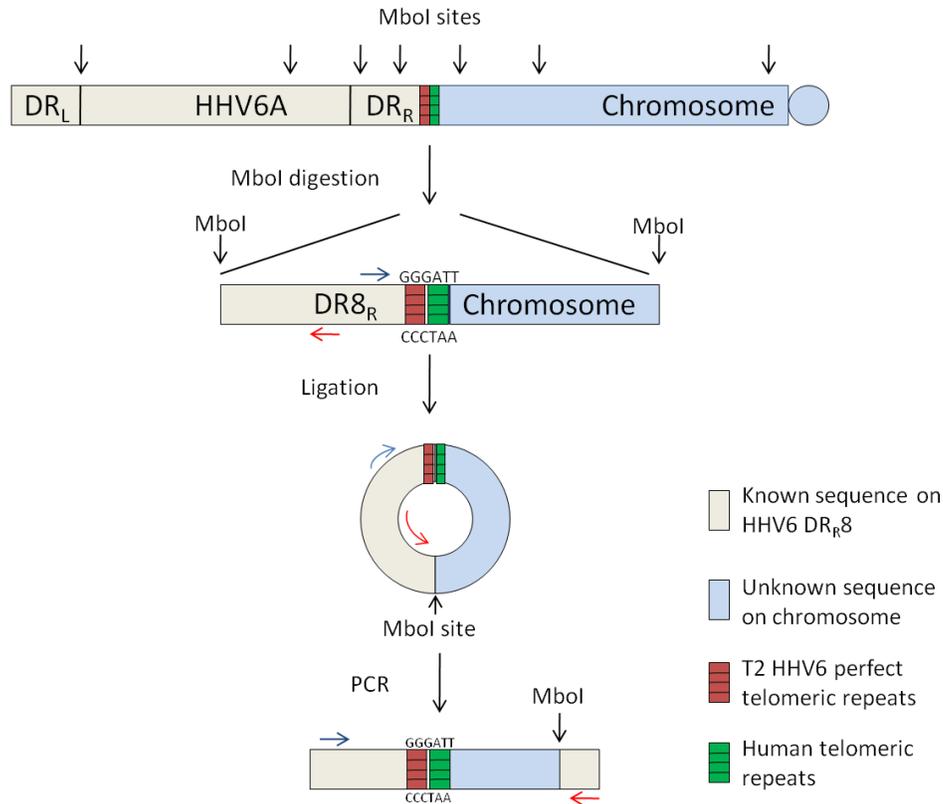


Figure 2.3. IPCR. Genomic DNA was digested with the restriction endonuclease *Mbol*, followed by ligation of the digested DNA at a very low concentration in order to promote the formation of circles. Primers anneal within the known DNA sequence of the DR8 of HHV6 and are orientated in opposite directions allowing the amplification of the full length of the circle as a linear fragment. The PCR product is sequenced using the same pair of primers or nested primers and the unknown DNA sequence from the human chromosome can be obtained.

Genomic DNA (1 μ g) was digested in a 50 μ l reaction with 10U of the restriction enzyme *Mbol* (NEB) at 37°C overnight. Digestion was stopped by heating at 65°C for 10 minutes. After complete digestion the DNA was diluted with D-H₂O and T4 ligation buffer (NEB) to a final concentration of 2ng/ μ l, 400U of T4 ligase (NEB) were added and ligation was left to proceed overnight at 16°C. Ligation was done with the digested DNA at low concentration to promote circularization of the DNA fragments. The ligation reaction was stopped by heating it to 65°C for 10 minutes. DNA was purified by extraction with equal volume of a phenol:chloroform:isoamyl alcohol and precipitated with ethanol (see Human genomic

DNA extraction from cell lines section). The ligated DNA was resuspended in 30µl of 10 mM of Tris pH 8.0 and a PCR was prepared using 1µl of the ligated DNA, 500 nM of each primer, 1X PCR buffer, 1.4U Taq polymerase, 0.3U of Pfu polymerase (Agilent technologies), in a final volume of 20 µl. The PCR programme consisted of an initial denaturing step at 94°C for 5 min, followed by 30 cycles at 94°C for 1 min, 64°C for 1 min, 68°C for 4min.

2.12.1 Southern blot preparation and hybridization to detect the chromosome-HHV6 junction fragment

In order to determine if the IPCRs contained the HHV6-Chromosome junction Southern blot analysis was carried out. Each completed PCR reaction (10 µl) was size separated on a 10 cm, 1% HGT agarose gel (Sakem) in 0.5X TBE at constant 100 V, using 1 Kb ladder as size a marker. A PCR product of 200 bp amplified from the DR8 region of the HHV6 genome was ³²PαdCTP labelled as a probe to detect the HHV6-Chromosome junction. Primers DR8F and DR8R used to generate this probe are shown in Table 2.5. The DR8-HHV6 probe was labelled and recovered as previously described for the telomere probe. The gel was blotted and hybridized as described in the Southern blot section above. Once the image with the DR8-HHV6 probe was obtained, the probe was removed by placing the membrane in a boiling solution of 0.1X SSC for 30 minutes with shaking. Afterwards the membrane was re-hybridised to the telomere probe.

2.12.2 Sequencing of the IPCR product

To sequence the IPCR product, 10µl of the PCR reaction were size separated on a 10cm, 1% HGT agarose gel in 0.5X TBE at constant 100 V, using the 1 Kb ladder as a size marker. As mention above, the IPCR products were not visible on agarose gels, therefore the gel was cut blindly according to the expected size of the band determined by the Southern

blot analysis. The IPCR product contained in the blindly cut agarose gel section was purified using the Zymoclean kit (ZYMO research) according to manufacturer's protocol. A second round of PCR with IPCR-1 and IPCR-2 primers was done to increase the amount of the amplicon. Amplicons from the second round PCR were size separated and purified from the gel as described above, then the purified amplicon was sequenced using Big Dye terminator V3.1 as described above.

2.13 Telomere Restriction Fragment (TRF) length analysis

Telomere Restriction Fragment (TRF) length analysis was conducted using pulsed field gels to measure telomere length in ALT+ cells because the heterogeneity of telomere length present in these cells.

Genomic DNA (5µg) was digested with 20 U of either *EcoRI* or a double digestion with *HinfI/RsaI* 20U of each at 37°C for 5hrs in a final volume of 30 µl. Digestion was stopped by heating the sample at 65°C for 10 minutes. To confirm the digestion was complete 300ng of each digest was size separated in a 10cm 1% agarose gel in 0.5X TBE run at constant 100 volts. In case of incomplete digestion 20U more of the enzyme were added and incubated at 37°C for 2hrs.

A 1% HGT agarose gel was prepared in 0.5X TBE without ethidium bromide and 4µg of the digested DNA were loaded into the gel along with 100ng of Lambda DNA-HindIII digest (NEB) and 100ng of Lambda monocut ladder (NEB) as size markers. Settings for the CHEF system were: initial time 0.5 sec, final time 1.5 sec, run time 14hrs at 6 Volts/cm, angle 120°. The gel was run in 0.5X TBE.

After the electrophoresis was completed the gel was stained with Ethidium bromide for 30 minutes with shaking to visualize the DNA. A photograph was obtained using a UV transilluminator. The gel was exposed to UV light for 5 minutes to generate nicks on the DNA strands and facilitate its transfer from the gel to the nylon membrane (exposure to

UV light shows better transfer in these gels, compared to depurination solution, own experience), and washed with denaturing and neutralizing solutions and Southern blot was prepared as described above.

Telomere length analysis was done using the Telometric software (<http://bioinformatics.fccc.edu/>).

2.14 Transfection using pSUPERIOR vector

2.14.1 Vector

To generate stable transfectants the pSuper RNAi SystemTM, from Oligo-Engine was used. The plasmid used was pSUPERIOR.neo+GFP. This vector uses the polymerase-III H1-RNA gene promoter, to produces a small RNA transcript lacking a polyadenosine tail and has a well defined start of transcription and termination signal consisting of five thymidines in a row (T5). The vector also contains a neomycin resistance cassette for selection of human cells with G418 after transfection and a GFP gene that can be detected in cells expressing the construct. A diagram of the vector is shown in Appendix 2.

2.14.2 Design of short hairpin oligonucleotides

Oligonucleotides that contain a unique 19 nucleotide (nt) sequence derived from the mRNA transcript of the target gene were designed. This sequence corresponds to the sense strand of the pSUPERIOR-generated siRNA. With this mechanism, the antisense strand of the siRNA duplex hybridizes to this region of the target mRNA to promote its cleavage and degradation. See figure 2.4 for an illustration of this process.

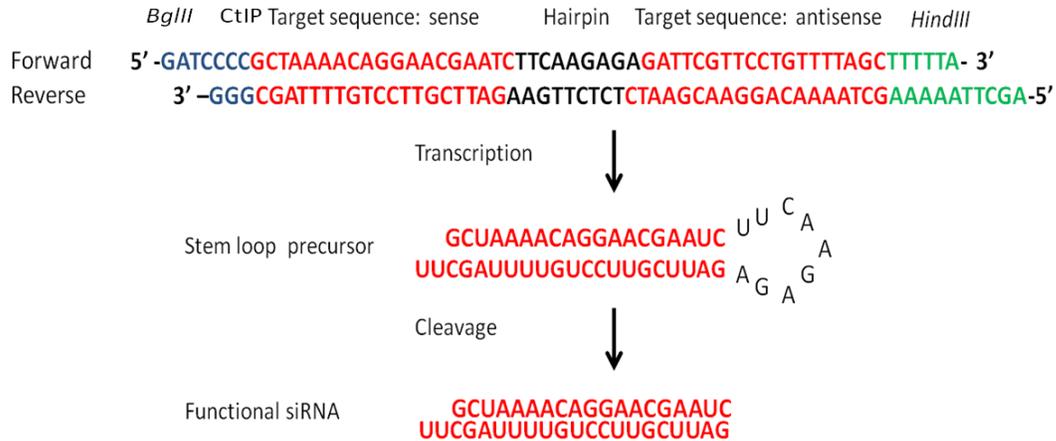


Fig 2.4 Diagram showing the annealed pair of oligonucleotides used to knock down the CtIP mRNA. The forward oligonucleotide contains the sense and antisense sequence of the target gene at its 5' end and 3' end respectively, after transcription the stem loop precursor is formed and then it is processed to generate the functional siRNA. The antisense sequence of the duplex siRNA generated hybridizes with the mRNA of the target gene promoting its cleavage and degradation.

According to the manufacturer's recommendations the oligonucleotide containing the sense sequence of the target mRNA at the 5' end was called forward and the complementary oligonucleotide was called reverse. The forward oligonucleotide contains the unique 19nt sequence in both sense and anti sense orientation, at the 5' end and 3' ends respectively, separated by a 9nt spacer sequence, which allows the formation of the hairpin. In order to be cloned into the plasmid, annealed oligonucleotides should have a *Bgl*III site at one end, and a *Hind*III site at the other end, which corresponds to the 5' and 3' end of the forward oligonucleotide respectively. By this means the forward oligonucleotide will be cloned downstream of the H1 promoter and the desired shRNA can be generated. Upon insertion of the annealed oligonucleotides into the pSUPERIOR vector the *Bgl*III site is destroyed which facilitated the selection of the recombinant vector containing the insert.

Two pairs of oligonucleotides were designed against each gene, EXO1 A and B and CtIP A and B. The sequence of EXO1 A and CtIP A were taken from publications (Bolderson E *et al.* 2009) (Huertas P *et al.* 2009), and the B versions of both were generated using the

“siRNA target finder tool” from Ambion (www.ambion.com). The sequence of the oligonucleotides is shown in table 2.11.

Table 2.11 Sequence of the oligonucleotides used to generate the shRNA against *CtIP* and *EXO1*.

ShCtIP-A	
A-1	GATCCCCGCTAAACAGGAACGAATCTTCAAGAGAGATTCGTTCTGTTTTAGCTTTTTA
A-2	GGGCGATTTTGTCTTGCTTAGAAGTTCTCTCTAAGCAAGGACAAAATCGAAAAATTCGA
ShCtIP-B	
B-1	GATCCCCTCCACAACATAATCCTAATTTCAAGAGAATTAGGATTATGTTGTGGATTTTTA
B-2	GGGAGGTGTTGTATTAGGATTAAGTTCTCTTAATCCTAATACAACACCTAAAAATTCGA
shEXO1-A	
A-1	GATCCCCTGTTGATCCTGAAACACTATTCAAGAGATAGTGTTCAGGATCAACATTTTTA
A-2	GGGACAACTAGGACTTTGTGATAAGTTCTCTATCACAAAGTCCTAGTTGTAAAAATTCGA
shEXO1-B	
B-1	GATCCCCGTCAGCTAATGTTAGCAGCTTCAAGAGAGCTGCTAACATTAGCTGACTTTTTA
B-2	GGGCAGTCGATTACAATCGTCGAAGTTCTCTCGACGATTGTAATCGACTGAAAAATTCGA

2.14.3 Vector cloning

The complementary oligonucleotides (encoding the shRNA) were annealed before being cloned into the pSUPERIOR plasmid. An annealing reaction containing 3µg of each primer and 48µl of annealing buffer (100 mM NaCl, 50 mM HEPES) was incubated at 94°C for 4 minutes in a hot-block, and then the reaction was left on the bench to cool down slowly for 1 hour. To check the annealing reaction an aliquot was loaded into a 4% NuSieve (Lonza) agarose gel in 0.5X TBE and size separated in 0.5X TBE at constant 100V. The annealed oligonucleotides were stored at -20°C until needed.

The pSUPERIOR plasmid (4µg) was linearized by digestion with 30U of *HindIII* (NEB) in 30µl

of 1X buffer 2 (NEB) for 1 hr at 37°C. After 1 hr, 30U of *Bgl*II were added and incubated for 90 minutes. The digestion reaction was stopped by heat inactivation at 65°C for 10 minutes. The digested plasmid was size fractionated in a 1% LE agarose gel in 0.5X TBE, and recovered from the gel using the gel extraction kit (QIAGEN). Then quantified with the Nanodrop 1000 (Thermo).

The linearized pSUPERIOR (500ng) was ligated to 2µl of the annealed oligonucleotides in 1X T4 DNA ligase buffer (NEB) and 400U of T4 DNA ligase. The reaction was incubated overnight at 16°C and then heat inactivated at 65°C 10 minutes. The ligated plasmids were treated with 10U of *Bgl*II at 37°C for 1 hr followed by inactivation at 65°C for 10 minutes. Using this digestion the plasmids without the insert are digested. Plasmids were loaded onto a 1% LE agarose gel in 0.5X TBE and two bands were observed the top band corresponding to the undigested plasmid and the bottom band corresponding to the digested plasmid, i.e. without the insert. The undigested plasmid was recovered from the gel with the gel extraction kit (QIAGEN).

2.14.4 Transformation of competent cells and plasmid purification.

Competent *E. coli* XL1-blue ampicillin and tetracycline sensitive were transformed with the ligated pSUPERIOR plasmids containing the inserts by electroporation. Gene Pulser *E. coli* Pulser™ Cuvettes 0.2cm (Bio-Rad) were placed on ice for 10 minutes. For each electroporation 40µl of *E. coli* (2×10^{10} cells/ml) were placed in ice cold tubes and 25ng of each plasmid were added. The bacteria and DNA mix was incubated for 5 minutes on ice and then transferred into the cuvettes and placed in the electroporation pod of the Gene Pulser Xcell™ electroporation system (Bio-Rad). Conditions of the electropulse given to the cells were capacitance 25µF, resistance 20ohms and voltage 250Vs, the pulse lasted for 4.5ms.

Luria broth (1ml) was added quickly to the cuvette and mixed. The bacteria were

transferred to a 15ml tube and incubated at 37°C with shaking (200 rpm) for 1 hr. After incubation the bacteria were plated (200µl/plate), onto Luria agar plates with ampicillin and tetracycline. The plates were incubated at 37°C overnight.

In order to identify the colonies harbouring the plasmid with the correct short hairpin insert, colony PCR was done using primers pSUP1 and pSUP2 which were provided by the supplier and are specific for the plasmid. Each PCR reaction contained 1X PCR buffer, 500nM of each primer, 0.5U of Taq polymerase and some cells from the colony being tested. Cycling conditions were 96°C for 5 minutes 1 cycle, 96°C for 20 sec, 65°C for 30 sec, 70°C for 60 sec for 26 cycles. PCR products were size separated in a 2.5% LE agarose gel in 0.5X TBE. The size of the expected product without the insert is 700 bp, and with the insert is 760 bp.

Colonies potentially harbouring the plasmid with the insert were re-plated on fresh agar plates and incubated at 37°C overnight. Afterwards one colony was taken from each plate and growth overnight at 37°C at 200rpm in 5ml of Luria medium with ampicillin and tetracycline. Plasmid DNA was extracted from 1.5ml of each culture using the E.Z.N.A. Plasmid Mini Kit I (OMEGA). Recovered plasmids were sequenced using Big Dye Terminator V3.1 reagent as described previously; each sequence reaction contained 500ng of plasmid DNA, primer pR5pSUPseq provided by the supplier (1.6 µM) 2µl, Big Dye Terminator Ready Reaction Mix 1µl, in 10 µl of 1X Big Dye Terminator Buffer. Cycling conditions were described in the PCR and sequencing section. After the sequence reaction the samples were cleaned as described before and sent to PNACL for electrophoresis. The electropherograms were analysed using the Autoassembler software.

In order to produce enough plasmid for Transfection, colonies carrying the correct insert were growth in 50ml of Luria medium with ampicillin and tetracycline, overnight at 37°C at 200rpm. 500µl of the culture of recombinant bacteria were stored in 500µl of 40% glycerol at -80°C as a stock. Plasmid DNA was purified from the bacteria using the Plasmid Midi Kit (QIAGEN). After measuring its concentration, the plasmid DNA was linearized by

digestion with *BamHI* (NEB). Each 50µl digestion reaction contained 15µg of DNA, 1X BSA, 20U *BamHI* in 1X buffer 3 (NEB). After digestion the plasmid DNA was purified with phenol/chloroform as described in DNA extraction section. The linearized plasmid was transfected into human cells.

2.14.5 Electroporation of human cell lines

Killing curves were conducted for SUSMI and WI-38 V13/2RA human cell lines to identify the capacitance and voltage at which 50% of cells remained alive. Cells were grown in the normal medium, as described in cell culture section, until they reached 80% confluence. The cells were then trypsinized and counted as described before, and then pelleted by centrifugation at 1100 rpm for 8 min at room temperature. The cells were resuspended in a volume of MEM medium to give one million cells in 0.8ml of medium. The cells (1×10^6) were transferred to a new 1.5 ml tube and incubated on ice for 5 minutes, then transferred to a 0.4cm Gene Pulser cuvette (Bio-Rad) and electroporated using the Gene Pulser Xcell™ electroporation system (Bio-Rad). The parameters used were capacitance at 950µF or 700µF combined with 200, 300, 400 or 500 volts and capacitance at 500µF combined with 200 or 500 volts. After electroporation cells were plated in 10 cm tissue culture grade Petri dishes in 15ml of MEM+10%FCS+1%NEAA medium and incubated as described in cell culture section. 24 hrs after electroporation the cells were trypsinized and counted in a Neubauer haemocytometer chamber. The kill curve experiment was done twice and the conditions giving ~50% survivals were used for further experiments.

2.14.6 Lipid-Base Transfection

To improve the transfection efficiency lipid based methods were evaluated for the SUSMI and WI-38 V13/2RA cell lines using the pSUPERIOR plasmid with no insert. The Table 2.12 describe the reagents tested and briefly the conditions in accordance with the

manufacturers' recommendation.

Table 2.12 Lipid based transfection. Brief description of the conditions used for each reagent tested for transfection.

Reagent	Reagent μ l	DNA μ g	Brief protocol according to supplier
jetPRIME (Polypus transfection)	4	1	Dilute DNA in jetPRIME buffer, add jetPRIME and incubate 10 min RT.
	6	1	
FuGENE (Promega)	9	2	Dilute DNA in MEM medium, add FuGENE and incubate 10 min RT.
	7	2	
Lipofectamine 2000 (Invitrogen)	6	2	Dilute DNA in Opti-MEM (Gibco), dilute Lipofectamine in same amount of Opti-MEM and incubate 5 min RT. Mix both dilutions and incubate 20 min RT.
	8	2	
Lipofectamine LTX (Invitrogen)	1.9	1.25	Dilute DNA in Opti-MEM, add Lipofectamine LTX, mix and incubate 30 min at RT.
	3	1.25	
Lipofectamine LTX/Plus reagent (Invitrogen)	1.9/0.6 Plus reagent	1.25	Dilute DNA in Opti-MEM, add Plus reagent and incubate 5 min RT, add Lipofectamine LTX, mix and incubate 30 min at RT.
	3/1.2 Plus reagent	1.25	

1.5×10^5 cells were plated in 6 well plates 24 hrs before transfection, so by the time of transfection the cells had reached ~70% confluence. Each mix described in the table containing DNA and transfection reagent was added to one well and cells were incubated as described in the cell culture section. After 6 hrs of incubation with the transfection reagent, the medium was replaced with fresh medium and 24 hrs after transfection G418 (500 μ g/ml) was added to select cells that contained the pSUPERIOR vector. The medium was changed every 72 hrs. After 14 days in selection the plates were fixed and stained with 2ml of 1% crystal violet solution (made in 70% Ethanol and 5% Formalin) for 10 minutes. The crystal violet solution was washed out with water and visible colonies were counted.

2.14.7 Colony forming assay to evaluate the effect of CtIP and EXO1 downregulation

The SUSM1 and WI-38 V13/2RA cell lines were transfected by electroporation using the best conditions obtained for each cell line by killing curves or by the best lipid-based method.

For electroporation 3 million cells were electroporated in duplicate reactions for each plasmid. Electroporation was carried out as described for killing curves, 4 μ g of each linearised plasmid DNA was added to the cells in the 1.5ml tubes and incubated for 5 minutes on ice. The cells from each electroporation were divided among three 10cm Petri dishes with 15ml of medium each. Twenty four hrs after electroporation the medium was changed and G418 (Gibco) at 500 μ g/ml was added for selection. The medium was changed every 72 hrs. After 15 days the culture medium was removed and the plates were rinsed with 5ml of 1X PBS. Cells were fixed/stained with 5ml of 1% crystal violet for 10 minutes. Excess crystal violet was removed with running water and colonies were counted. The experiment was done twice for each plasmid.

For the lipid-based method, one million cells were plated in 10cm petri dishes in 15ml of medium 24 hrs before transfection. On the day of transfection 14 μ g of each plasmid DNA were diluted in 1.5ml of MEM medium and mixed. The FuGENE reagent (50 μ l) was added to the same tube and the mix was incubated for 10 minutes at RT and before being added to the cultures. After 6 hrs the medium was replaced with fresh medium and 24 hrs after selection applied (G418 at 500 μ g/ml). As above, the medium containing G418 was changed every 72 hrs for 15 days. After 15 days in culture medium was removed and plates were rinsed with 5ml of 1X PBS, cells were fixed/stained with 5ml of 1% crystal violet for 10 minutes. Excess crystal violet was removed with running water and colonies were counted. The experiment was done twice for each plasmid.

2.14.8 Single-cell colony isolation and expansion

Colonies that grew from SUSM1 cells that had been transfected by electroporation, and WI-38 V13/2RA that had been transfected using FuGENE reagent (as described above) were identified. Once the colonies were ~0.5 cm in diameter, the plates were rinsed with 10 ml of 1X PBS and air dried for five minutes. Trypsin (10 μ l) was placed over each colony they were recovered by placing a six mm cloning disc (Scienceware[®] Sigma-Aldrich) soaked in trypsin on top of the colony for five minutes at RT. This was carefully removed with sterile tweezers to a six well plate with 2.5ml of the selection medium.

Successfully recovered clones were grown in six well plates until they reached 70% confluence, then trypsinised and reseeded in 25cm flasks. When they reached 80% confluence, cells were trypsinised, counted to calculate the number of population doublings, and split among five 25 cm flasks. The expanding clones were passaged to new flasks when they reached 80% confluence and the number of population doublings recorded at each passage.

2.14.9 Population doublings calculation

When the five flasks from one clone reached 80% confluence, the cells were trypsinised and pooled in one 50ml tube with 25ml of fresh medium. Cells were counted in a Neubauer chamber and populations doublings were calculated using the formula $PDL = (\log_{10} \text{ cell count at harvest} - \log_{10} \text{ cell count at inoculation}) / 0.301$. The value 0.301 is a constant for the conversion factor from the \log_{10} to the natural log (\log_2).

2.15 siRNA transfection

In addition to the transfection using the pSUPERIOR vector, Silencer[®] Select Predesigned siRNA (Ambion) against CtIP (sequence of the siRNA had been designed by the supplier

Appendix 1) was used to look for a rapid effect on C-circle formation in ALT+ cell lines. Based on previous evidence C-circles seem to be reliable markers of ALT activity (Henson JD *et al.* 2009). At the same time downregulation by siRNA allowed us to observe changes in the localization of CtIP foci by Immunofluorescence (IF).

The lyophilized siRNA (5nmole) was resuspended in 100 μ l of RNase free D-H₂O to get a stock at 50 μ M stored at -20°C to avoid degradation. From this stock a dilution to 10 μ M was made with RNase free D-H₂O and stored at -20°C. Two 10 cm petri dishes, each containing one million cells in 12ml of medium were prepared for each cell line 24hrs before transfection. A 2cm² coverslip was placed at the bottom of each petri dish to allow cells to grow on the surface of the coverslip. The coverslip provided a sample of the siRNA treated cells to be analysed by immuno-fluorescence. On the day of transfection the cells were ~80% confluent. According to the manufacturer's recommendations, 20 μ l of Lipofectamine 2000 (Invitrogen) were diluted in 1.5ml of OptiMEM (GIBCO) medium and incubated for five min at room temperature, in a different tube 30 μ l of the siRNA (10 μ M) were diluted. After the five minute incubation Lipofectamine 2000 and siRNA dilutions were mixed and incubated for 20 min at room temperature to allow the formation of the lipid-siRNA complexes. The mix was then added to the cells and the cultures incubated at 37°C.

After 48 hrs in culture, the media was removed and the cells were rinsed in 1X PBS, the coverslips were recovered from the bottom of the petri dishes and immuno-fluorescence was conducted as described above. The PBS was removed and remaining cells were trypsinised and harvested as previously described. The cells obtained from each petri dish were divided among two pellets one for DNA extraction to perform the C-circle assay and one for RNA extraction to analyse by qPCR the effect of the siRNA on the expression of the CtIP gene.

2.16 C-Circle assay

The C-Circle assay is a DNA amplification method to detect circular, partially single stranded, extrachromosomal telomeric DNA. The method was described by Henson JD *et al.*, (Henson JD *et al.* 2009), and takes advantage of the properties of the DNA polymerase from the phage phi 29 (Φ 29). The technique was described as a marker for ALT activity.

Genomic DNA was digested with *Eco*RI (as previously described), which should leave telomeric circular DNA intact. The digested DNA was diluted to 5ng/ μ l. Each amplification reaction contained Φ 29 DNA polymerase 7U (Fermentas LTD), 1X Φ 29 DNA polymerase buffer, BSA 200 μ g/ml, 0.1% Tween 20, DTT 1 μ M, dATP 1 μ M, dTTP 1 μ M, dGTP 1 μ M and 10ng of digested DNA, in a final volume of 10 μ l. The reaction was incubated at 30°C for eight hrs followed by inactivation of the Φ 29 DNA polymerase at 65°C for 20 min. Half the reaction was denatured by adding 5 volumes of denaturing solution (0.5 M NaOH, 2M NaCl, 25mM EDTA) to remove secondary structures, and dot-blotted on Hybond XL positively charged membrane (Amersham), using Dotblot loading dye (30% glycerol, 0.5X TBE, bromophenol blue). Each well was rinsed with 120 μ l of 2X SSC to neutralize the membrane, which was then dried at 80°C for 10 min and UV crosslinked. The membrane was hybridized, washed and developed as described in Southern blot section.

2.17 CtIP detection by Immunofluorescence and FISH detection of telomeric DNA using a Peptide Nucleic Acid probe

The cells were plated on 2cm² coverslips. After 24 hrs the medium was removed and the coverslips were rinsed in 1X PBS 4 X 5 min each. The cells on the coverslips were fixed and permeabilized with cold acetone (-20°C), for 20 minutes at -20°C, and then were rinsed with 1X PBS four times (5 min each) at RT. The cells were blocked in 3% BSA in 1X PBS for 1 hr at RT. The CtIP antibody (ABCAM ab70163) was diluted 1:1000 in 3% BSA in 1X PBS and cells were incubated for 1hr at RT. Coverslips were rinsed with 1X PBS four times (5 min

each) at RT. The secondary antibody (mouse anti-rabbit Alexa FluorR[®]488 Invitrogen) was diluted 1:1000 in 3% BSA in 1X PBS and cells were incubated 1 hr at RT. Cells were washed 4 four times (5 min each) at RT with 1X PBS. The cells were then fixed in 4% paraformaldehyde in 1X PBS for 2 min at RT. The coverslips were rinsed in 1X PBS twice for 5 min. Dehydration was with 70% ethanol, 80% ethanol and 100% ethanol 2 min in each at RT. The coverslips were then left to air dry. The PNA-telomere probe was a C rich 18mer oligo labeled with Cy3 (Panagene). The PNA-probe (50 μ M) was diluted 1:100 in hybridization mix (70% formamide, 0.25 blocking agent, 10mM Tris pH 7), 5 μ l of the PNA probe-hybridization mix were placed on a slide and the coverslip lowered on top avoiding bubbles. The coverslips on the slide were sealed with cement solution in the dark. The cells on the coverslips were denatured on an Omnislid at 80°C for 3 min and the left to hybridize at 22°C for 2hrs. Afterwards the rubber cement was removed and coverslips were carefully inverted for posthybridization washes in 70% formamide, 10mM Tris pH 7.2 with shaking (2X10 min), and then in 0.05MTris pH 7.5/0.15M NaCl/0.05% Tween 20 (2X5 min). Dehydration was done in 70% ethanol, 80% ethanol and 100% ethanol 1 min each at RT. Coverslips were mounted with a Vectashield and DAPI mix (10 μ l of Vectashield per 1 μ l of DAPI at 50 μ g/ml) and analyzed on an Olympus confocal FV1000 microscope.

2.18 MS32 analysis

Analysis of the minisatellite MS32 located in chromosome 1 was performed, as described previously (Jeyapalan JN *et al.* 2005). Genomic DNA (1 μ g) was fully digested with *Mbo*I restriction endonuclease at 37°C, which does not cleave the MS32 repeat array. The DNA concentration was calculated using the Nanodrop1000 (Thermo) and further diluted to the required concentration with high molecular weight salmon sperm DNA as a carrier. Primers MS32B and MS32E (see table 2.13) amplify all alleles and anneal 380 and 75 bp upstream and downstream from the repeat array respectively (Jeyapalan JN *et al.* 2005). Each PCR reaction contained 1X PCR buffer, 300nM of each primer, Taq polymerase/Pfu at

a ratio of 20:1 (0.1U and 0.005U respectively), digested DNA (6pg-1ng), Tris pH8 12.5mM and high molecular weight salmon sperm DNA (as a carrier) at 1ng/ μ l in a final volume of 10 μ l. Cycling conditions included an initial denaturing step at 96°C for 1min, followed by 22 cycles at 96°C for 20 sec, 62°C for 30 sec, 68°C 6min, and final extension step at 65°C for 7 min.

Table 2.13 Primers used for amplification of the MS32 minisatellite. Obtained from (Jeyapalan JN *et al.* 2005)

Primer	Sequence 5'→3'	Length bp	TM °C
MS32B	AAGCTCTCCATTTCCAGTTTCT	22	62
MS32E	CTTCCTCGTTCTCCTCAGCCCTAG	24	

PCR products were size separated in 0.8% LE agarose gels in 0.5X TBE, and Southern blot hybridization was carried out as above. Membranes were hybridized with a ³²P-labelled MS32 tandem repeat probe.

In order to quantify the number of mutant molecules, and obtain mutation frequencies, Small Pool PCR (SP-PCR) was conducted using approximately 20 amplifiable molecules per reaction. The number of amplifiable molecules was calculated with Poisson analysis as described previously (Jeffreys AJ *et al.* 1994).

To calculate the number of amplifiable molecules per reaction PCR reactions containing an estimated 3pg, 6pg or 12pg of digested DNA were prepared, 6pg of genomic DNA were considered as the equivalent of a diploid genome. For each DNA amount 10 reactions were prepared, based on 50% of PCR efficiency the amount of DNA at which half of the reactions showed at least one PCR product, was considered as the amount of DNA equivalent to 1 amplifiable molecule. Forty PCR reactions using DNA equivalent to 1 amplifiable molecule were prepared alongside the SM-PCR, the number of positive reactions (expected around 50%) was used to calculate the number of molecules on each SM-PCR and to obtain the rate between normal and mutant alleles. The rate was obtained

by dividing the total number of molecules in the SM-PCR by the number of mutants counted (total number of molecules/number of mutant allele).

2.19 ss telomeric DNA assay

To analyse the 3' overhang present at telomeres a method previously described (Zhao Y *et al.* 2007) was used. This method takes the advantage of the double stranded specific endonuclease (DSN) extracted from the Kamchatka crab (Evrogen, Russia). The enzyme is able to degrade all double stranded DNA to fragments ~10bp, leaving just single stranded fragments. Genomic DNA (5µg) was digested with the 0.2U of DSN at 65°C for 2 hours and then 1 hour at 37°C. The enzyme was inactivated with its own stop buffer supplied by the manufacturer. As a negative control 5µg of the same DNA was digested with 10U of *E. coli* Exonuclease I at 37°C for 1 hour, which degrades overhangs in 3' to 5' direction. This enzymatic activity should degrade the telomeric 3' single strand overhangs. After digestion with Exonuclease I the DNA was extracted and precipitated as described in DNA extraction section, before treated with the DSN. This purification/precipitation was needed because the Exonuclease I buffer interferes with the activity of DSN.

After DSN digestion, samples and negative controls were analyzed by dot blot. Samples were denatured by adding five volumes of denaturing mix (0.5 M NaOH, 2M NaCl, 25mM EDTA), and dot-blotted on Hybond XL positive charged membrane (Amersham), using the dot blot loading dye (30% glycerol, 0.5X TBE, bromophenol blue). Each well was rinsed with 120µl of 2X SSC to neutralize the membrane; the membrane was dried at 80°C for 10min and UV crosslinked. The dot blot membrane was hybridized to the telomere probe as described in the southern blot section. Images were analyzed using the Image J dot blot analyzer plug in (<http://rsbweb.nih.gov/ij/>).

DSN treated genomic DNA samples were also analysed in denaturing polyacrylamide gels, as described by Zhao Y *et al.* After stopping the DSN digestion reaction, the same volume

of formamide loading dye was added and the samples heated at 65°C for 5 min. Samples were loaded in 6% denaturing polyacrylamide gel containing 8M urea. Electrophoresis was conducted at constant 30W for 30 min with Low Molecular Weight DNA ladder (NEB) as size marker. DNA was transferred to a Hybond XL positive charged membrane by Southern blot over night then the membrane was dried at 80°C for 10 min and UV crosslinked. Membrane was hybridized to the telomere and respective ladder probes as described in the Southern blot section.

CHAPTER 3 – ROLE OF THE WRN AND BLM HELICASES IN THE ALT PATHWAY

3.1 Background

Current evidence has shown the functional and physical interaction of WRN and BLM with at least three members of the shelterin complex, namely TRF1, TRF2 and POT1 (Opresko PL *et al.* 2002) (Opresko PL *et al.* 2005). The participation of these helicases in the ALT pathway is supported by the presence of both proteins in the APB's (Nittis T *et al.* 2008). BLM and TRF2 can colocalize within APBs and this colocalization seems to correlate with BrdU incorporation and it is increased during late S and G2 phases when ALT is thought to occur. In addition BLM is also able to colocalize with TRF1 in the APBs and co-immunoprecipitation experiments demonstrated that this physical interaction is not DNA-dependent (Nittis T *et al.* 2008). Interestingly TRF1 and TRF2 modify BLM's capacity to unwind telomeric substrates in opposite ways. While TRF1 inhibits the helicase activity on telomeric substrates, TRF2 is able to enhance it on telomeric and non-telomeric substrates, suggesting that the control of the BLM helicase activity at telomeres is dependent on the concentration of TRF1 and TRF2 (Lillard-Wetherell K *et al.* 2004). More evidence supporting the importance of BLM in telomere maintenance in ALT cells was provided by Temime-Smaali *et al.*, based on the knowledge that BLM interacts with Topoisomerase III (Topo III), RMI1 and RMI2 to form a complex called RTR (RecQ/Topo III/RMI). The main function of this complex is to resolve recombination intermediates, which is essential for genomic stability and they showed that Topo III can form a complex with TRF2 and BLM in ALT cells. It has also been shown that Topo III is an important factor in ALT cells since reduced expression by siRNA results in increased cell death in ALT cells but not in telomerase positive cells. Furthermore inhibition of Topo III expression in ALT cells leads to telomere uncapping and reduced levels of TRF2 and BLM (Temime-Smaali N *et al.* 2008). These authors also showed that treatment of ALT cells with the stabilizer of G-cuadruplexes (telomestatin) leads to telomere uncapping. This is shown by a massive

increase in DNA damage response (DDR) signalling through γ H2AX accumulation, accompanied by depletion of the Topo III/BLM/TRF2 complex. These results point to the importance of the RTR complex in the resolution of aberrant structures at telomeres in order to conserve telomere function (Temime-Smaali N *et al.* 2009).

In 2009 Bhattacharyya *et al.*, using double immunoprecipitation and mass spectrometry reported three novel partners of BLM in ALT cells. These were telomerase protein 1 (TEP1), heat shock protein 90 (HSP90) and topoisomerase II α (TOPO II α). These proteins associate with BLM and TFR2 in APBs *in vivo*. These associations were observed at sites of new DNA synthesis as suggested by the incorporation of BrdU, and it was shown that the association was not dependent on DNA. The authors also identified another complex formed by TRF1, TFR2, BLM, TOPO II α and HSP90, but without PET1. Functional studies revealed that HSP90 inhibits the enzymatic function of BLM on telomeric and non-telomeric substrates, whereas TEP1 and TOPO II α , inhibit BLM unwinding activity on telomeric substrates just at the beginning of the reaction. In the same study the authors showed by telomere restriction fragment (TRF) length analysis a significant increase in the erosion rate of telomeres after the downregulation of BLM. This effect was not seen in telomerase positive cells, and there was no formation of the complexes with TRF2 upon BLM silencing. These findings suggest the existence of dynamic BLM-associated protein complexes in ALT cells that can participate in regulating the enzymatic functions of BLM during telomere elongation (Bhattacharyya S *et al.* 2009).

Post-translational modifications also seem to play a role in the function of BLM. It has been reported that the sumoylation of the yeast homolog of BLM and WRN, Sgs1, on the serine 621 favours the development of Type II survivors. The exact mechanism for this effect is not understood yet, interestingly WRN and BLM are also subject to sumoylation. Whether this modification is more abundant in ALT cells compared to telomerase positive cells remains to be investigated (Lu CY *et al.* 2010). Experiments in telomerase negative yeast strains with deletion of the *Sgs1* gene showed that BLM can complement the activity of Sgs1 and induces the proliferation of Type II survivors (Lillard-Wetherell K *et al.* 2005).

The Accumulating evidence has suggested that BLM is an important participant in ALT cells, however the exact role in the pathway remains to be clarified.

The recombinase RAD52, which is essential for the generation of survivors in the absence of telomerase in yeast, has been shown to be an interacting partner of WRN. Fluorescence resonance energy transfer (FRET) experiments demonstrated that WRN and RAD52 interact *in vivo* at stalled replication forks. In addition, *in vitro* analysis showed that RAD52 is able to regulate WRN helicase activity in a substrate-dependent manner. Finally it has been demonstrated that WRN can stimulate RAD52 strand annealing activity between single stranded DNA and a homologous sequence contained within a duplex DNA molecule (Baynton K *et al.* 2003). Considering that both proteins can be located in APBs, these findings suggest a collaborative role between RAD52 and WRN that may be important for telomere replication in ALT cells. Another protein which functionally interacts with WRN is the Flap endonuclease 1 (FEN1), a structure specific nuclease important for the processing of Okazaki fragments during replication. FRET analysis supports the interaction of these proteins *in vivo* on stalled replication forks. Also *in vitro* experiments showed that WRN is able to stimulate the nuclease activity of FEN1 (Sharma S *et al.* 2004). Recent studies have demonstrated that downregulation of *FEN1* expression in ALT cells results in an increase in the frequency of Telomere induced DNA damage foci (TIFs) as well as end to end fusions. However these effects were not observed in telomerase positive cells. These data indicate that co-operation between WRN and FEN1 is important for resolving complex structures that arise during telomere replication in ALT cells (Saharia A & Stewart SA 2009). Despite the evidence of WRN being associated with key genes for the ALT pathway, it does not seem to be indispensable for ALT cells, as telomerase null mouse embryo fibroblast (MEFs) lacking WRN show telomere dysfunction and a high rate of T-SCE indicative of increased recombination within telomeric DNA. These *TERC*^{-/-} *WRN*^{-/-} cells escape from senescence and emerge as immortalized clones that engage the ALT mechanism (Laud *et al.* 2005).

Despite all the accumulated evidence that WRN and BLM may play roles in ALT the

definitive role for each enzyme is not fully understood. The analysis of WRN and BLM helicases was started by Dr. Mendez-Bermudez in Dr. Royle's laboratory. He used the W-V cell line, which is a cell line derived from an SV-40 immortalised fibroblasts from a patient with WS. The cell line uses the ALT pathway as a telomere maintenance mechanism (TMM). At the initial stages of the study he showed a reduction of colony forming efficiency of the W-V cell line following transfection with a shRNA against the *BLM* helicase (referred as shBLM hereafter), compared with the W-V cell line transfected with an empty vector. Dr. Mendez-Bermudez also generated clones from the W-V cell line with downregulation of *BLM* expression, by Western Blot, compared to W-V control clones transfected with an empty vector. Among the W-V shBLM clones, clone 6 showed the highest level of BLM downregulation (~60%). The W-V shBLM clone 6 was used for analysis of the MS32 minisatellite and telomere mutation frequency. I participated in the telomere mutation analysis on the W-V cell line, W-V shBLM clone 6 and W-V control clone 4. I want to acknowledge Dr. Mendez-Bermudez and Dr. Cotton for initiating this study and their guidance.

3.2 Aim

To study the role of the WRN and BLM helicases in the ALT mechanism by analysing mutation frequency and profile in an ALT + cell line that lacks the WRN protein and then following depletion of BLM through the expression of a shRNA, in a clone derived from the WRN^{-/-} ALT+ cell line.

3.3 Analysis of telomere mutations in the ALT+ cell line WV

The W-V cell line was generated from a patient with WS, which means the cell line lacks the WRN protein. The W-V cell line was immortalized with the SV-40 large T antigen and does not show telomerase activity, it maintains telomere length via the ALT pathway. To

investigate the role of the WRN helicase in the ALT pathway the mutation frequency was determined in the W-V cell line. In order to achieve this, single telomere molecules were amplified by STELA using allele specific primers for the XpYp subtelomeric region with Telorette2 and teltail primers (Figure 3.1A shows an example of a Southern Blot containing single STELA products from the W-V cell line used for the analysis). From the single STELA products, Telomere Variant Repeats maps were obtained by TVR-PCR analysis using a nested subtelomeric primer and primers for the telomere repeat (TTAGGG/TagTel W) and variants (for example TGAGGG/TagTelX) (Figure 3.1B is an example of one of the gels used for TVR mapping of the single STELA products from W-V clone 6). A total of 474 maps from STELA products from the W-V cell line were analysed and no mutations were seen in 468 molecules (i.e. TVR maps from these molecules were identical to progenitor allele depicted in figure 3.2A). However six mutant molecules were identified (Figure 3.2A) and each was found just once, this represents a mutation frequency of 1.3% (Table 3.1) (Dr. Mendez-Bermudez *et al*, paper submitted).

Previous mutation analysis by TVR mapping in ALT cell lines revealed a mutation frequency of around 5.5%. Among these mutants, approximately 70% were characterized as complex mutations. A complex mutation implies a truncation of the progenitor allele and addition of repeats from a non sister-chromatid molecule that could be derived from another telomere or extrachromosomal telomeric DNA. The result is a map unrelated to the progenitor allele distal to the breakpoint (Varley H *et al*. 2002). The mutant molecules observed in the W-V cell line (Figure 3.2A) are different from the complex mutations previously seen in other ALT cell lines. For example in four mutants (1, 2, 3 and 5) there was a truncation of the progenitor allele and substitution of the distal TGAGGG repeats with a pure array of TTAGGG repeats. In addition, mutant number one also showed a loss of one TGAGGG repeat from the initial block. Mutant number four showed an insertion of one TTAGGG in the initial block of TGAGGG and in mutant number six there was an insertion of eight TTAGGG repeats interrupted by one TGAGGG (Figure 3.2A). The lack of “complex mutations” in the W-V cell line suggests that WRN participates in the generation

of such mutations commonly observed in other ALT cell lines.

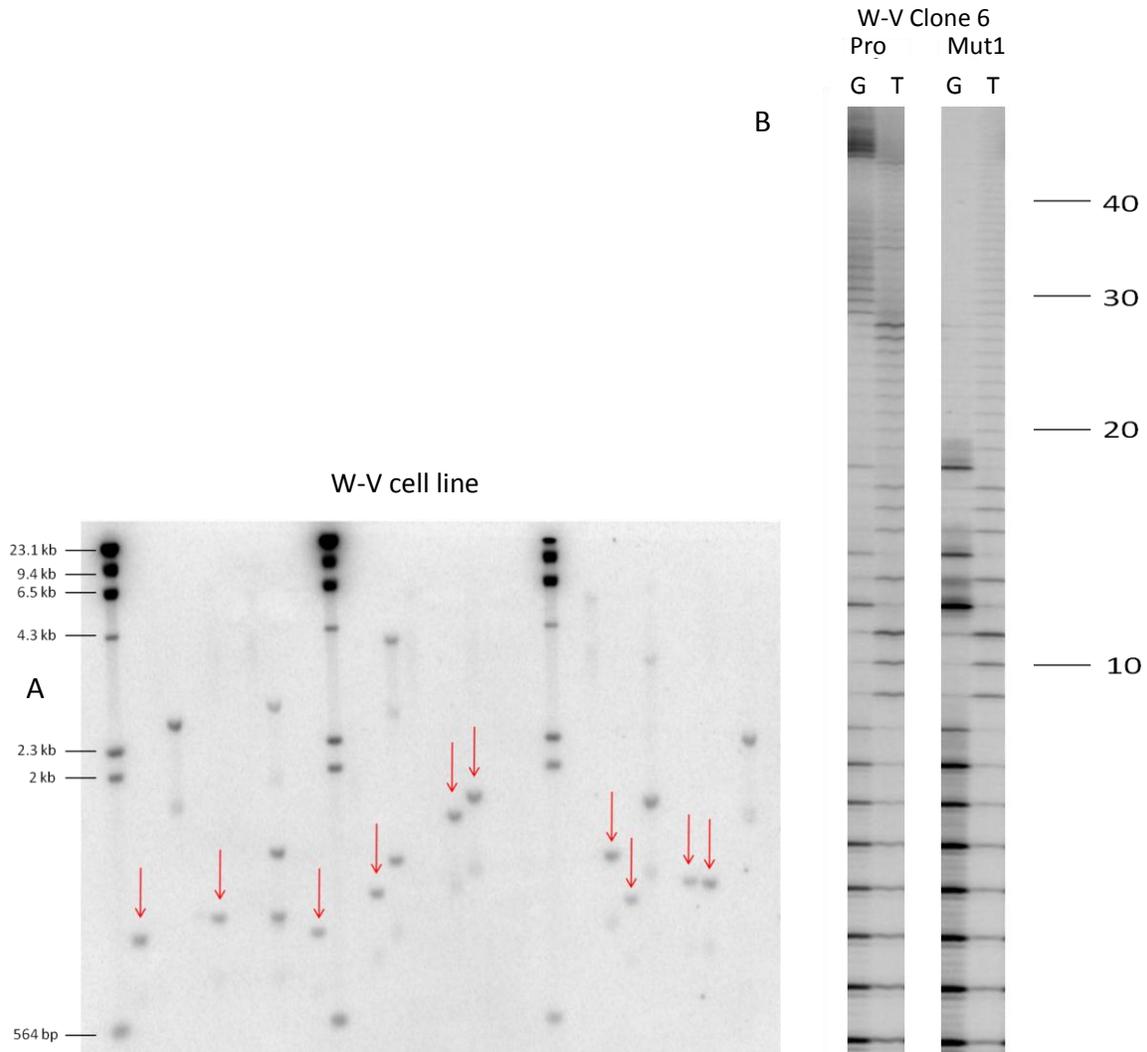


Figure 3.1 XpYp STELA products and TVR maps from the W-V cell line and clone 6. A) Example of one STELA conducted on small aliquots of W-V cell line genomic DNA. Each 10 μ l reaction contained approximately one amplifiable molecule. The products were detected, following agarose gel size fractionation, by Southern blot hybridisation to the 32 P radiolabelled telomere repeat probe. Arrows point to single molecules (i.e. single telomeres) selected for further analysis by TVR-PCR to determine interspersions of TTAGGG and variant repeats. B) Example of TVR maps from single telomere (STELA) products from W-V shBLM clone 6. "Pro" indicates the progenitor, while Mut1 represents the mutant number one from the W-V clone 6 (see figure 3.2C). In both examples the left lane corresponds to the TGAGGG (G) repeats, while the right lane corresponds to the TTAGGG (T) repeats. Note that in the mutant allele the distal array of G repeats is substituted by T repeats. Numbers on the right indicate the repeat number from the start of the telomere.

3.4 TVR mapping on clones with downregulation of *BLM*

To investigate the function of the *BLM* helicase in the ALT+ W-V cell line, it was transfected with the pSUPERIOR vector containing a shRNA against *BLM*, to obtain clones with reduced levels of *BLM* expression. The cells were transfected by electroporation with the plasmid containing the shRNA against *BLM* or an empty vector as a control. Visible clones were recovered and expanded and after ~21PD (population doublings) each clone was tested for *BLM* expression by Western blot. Most of the clones showed less than 20% downregulation compared to empty vector, but one of them, clone 6, had ~60% downregulation (data not shown). This work was conducted by Dr. Mendez-Bermudez (Mendez-Bermudez paper submitted).

To investigate the effect of *BLM* downregulation on the telomere W-V mutation rates and mutation profile, single molecule STELA and TVR-PCR were conducted. Genomic DNA from W-V shBLM clone 6 and a W-V control clone 4 (transfected with an empty vector) was extracted and digested with *EcoRI*. STELA was carried out to obtain single telomere molecules using an allele specific primer for the XpYp subtelomeric region and mapping was done by TVR-PCR. In total 216 molecules from W-V shBLM clone 6 were analysed and 30 mutant molecules were found (of eight different mutant types). This corresponds to a mutation rate of 6.9×10^{-3} per PD (Table 3.1 and Figure 3.1). In the W-V control clone 4, a total of 110 molecules were analysed revealing three mutant molecules, representing a mutation rate of 1.4×10^{-3} per PD (Table 3.1). These data revealed an increase in mutation frequency upon *BLM* downregulation. The interpretation of the mutation rate in W-V shBLM clone 6 should be done cautiously because six of the eight mutant types were seen more than once, for example the mutant 1 was seen 14 times in total. The possible explanations for this are that each mutant arose as an independent event, that some mutant molecules (e.g. W-V shBLM clone 6, mutant 1) originated early during the clonal expansion so giving rise to somatic mosaicism, or a combination of both. The possibility of contamination of the STELA PCR reaction was ruled out because all the STELA products

from W-V shBLM clone 6 and control clone 4 used for TVR mapping have different lengths.

The fact that the map from mutant 1 of the W-V shBLM clone 6 was also observed in mutant 2 from the progenitor W-V cell line favours the idea of each mutant being the result of independent events. The same is true for mutants 2 and 3 of the W-V shBLM clone 6, which were seen more than once in the W-V shBLM clone 6. The map of W-V shBLM mutant 2 is identical to the map of the mutant 3 of the progenitor W-V cell line, and mutant 3 from the W-V shBLM clone 6 was seen in the mutant 3 of the W-V control clone 4. In summary the data obtained showed an increase in mutation rate after *BLM* downregulation in the W-V cell line. The mutations observed are not complex mutations but they are similar to those observed in the progenitor W-V cell line.

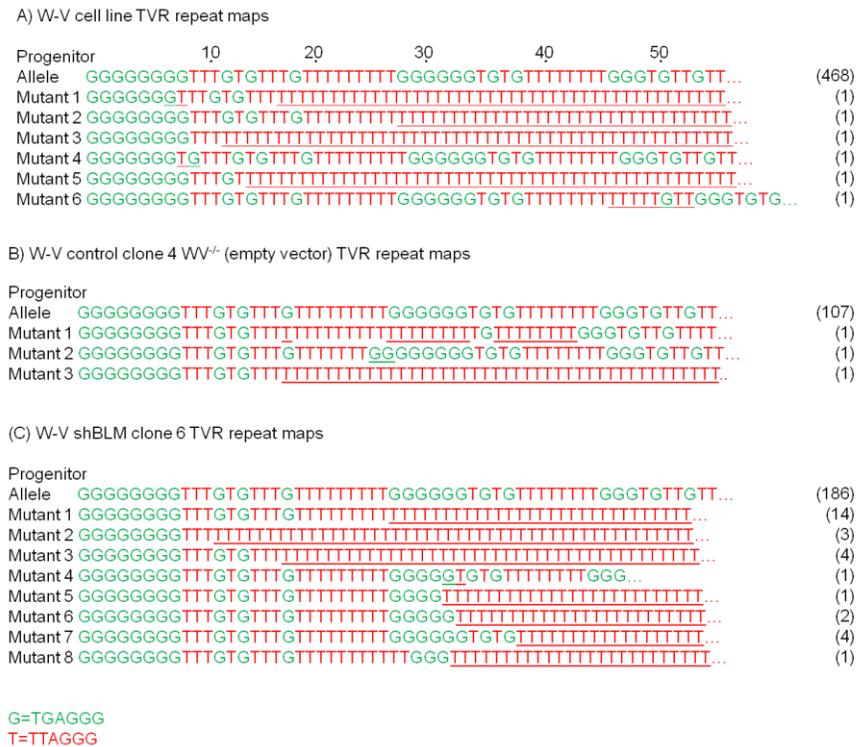


Figure 3.2 Telomere Variant Repeats maps from the progenitor allele of the XpYp telomere in the WV cell line (A) and control and BLM depleted clone (B and C respectively). Mutations found in each analysed cell line or clone are shown. The number on the right represents the number of times that a particular mutant was seen in different STELA products. The underlined letters depict the site where the pattern of repeats changes. The numbers across the top of the progenitor allele in A represent the number of repeat in the telomere.

Table 3.1 Mutation frequencies and rates for the groups studied. ND not determined

Group	Molecules analysed	Mutants detected	Mutation frequency	Mutation rate
W-V cell line (W-V ^{-/-}) *	474	6	1.3%	ND
W-V control clone 4 *	110	3	2.7%	1.4x10 ⁻³
W-V shBLM clone 6 (40% BLM expression)	216	30 (8 different)	13.8%	6.9x10 ⁻³

*Molecules analysed in collaboration with Dr. Cotton

3.5 Discussion

The W-V cell line uses the ALT pathway for telomere maintenance and lacks WRN expression and so this cell line provides a model to study the effect of the absence of WRN on the ALT mechanism. The W-V cell line showed a telomere mutation frequency of 1.3% (6/474) and the profile of mutants observed can be explained by intra-allelic events such as unequal telomere sister-chromatid exchange, although participation of a non-homologous telomeres or extrachromosomal DNA cannot be completely ruled out. Interestingly there was no evidence of mutants containing “complex mutations”. In previous reports the average mutation frequency in ALT cells was 5.5% and ~70% of the mutations were characterized as complex mutations (Varley H *et al.* 2002). This finding strongly suggests that the generation of such complex events is related to WRN activity and/or its interaction with proteins involved in the ALT pathway, such as RAD52. RAD52 seems to play an important role in the telomerase-independent maintenance of telomeres and its function can be stimulated by WRN (Baynton K *et al.* 2003). Transfection of W-V cells with vectors expressing wild type WRN or a helicase-deficient or a nuclease deficient version of the enzyme would help to have a better understanding of the mechanism underlying the lack of complex mutations. Unfortunately attempts to express WRN in the W-V cell line were not successful (Mendez-Bermudez *et al* paper submitted).

When the expression of the BLM helicase was downregulated by 60% in the W-V shBLM clone 6, the telomere mutation frequency rose to 13.8% giving a mutation rate of 6.9×10^{-3} per PD whereas in the W-V control clone 4, transfected with an empty vector, the mutation frequency was 2.7% and mutation rate 1.4×10^{-3} per PD respectively. The mutation rates between the BLM depleted and control clones were found to be statistically different (Fisher's exact test, two tailed. $p < 0.005$). However because six of the eight different mutations were seen more than once, the presence of somatic mosaicism is possible. This scenario would reduce the mutation frequency, if we consider just the eight different mutants observed in the clone 6, then the mutation frequency is 3.7% (8/216), which is not significantly different from the W-V control clone 4. The presence of long tracts of the same repeat makes it difficult to know the exact breakpoint. For example, the progenitor molecule has an array of nine TTAGGG repeats followed by six TGAGGG repeats but in mutant 1 from the W-V shBLM clone 4 (seen 14 times), there was a deletion of these TGAGGG repeats and homogenisation to an array of pure TTAGGG repeats. The breakpoint could have happened on any of the nine TTAGGG repeats and produced the same map, but they would be mutations originated from independent events. The fact that the same mutation was observed in clone 6 and the progenitor W-V cell line, and that all the STELA products are different in length, supports the possibility that each mutant represents an independent event. In spite of an increase in the mutation rate, there was no evidence of "complex mutations", and all the mutants found can be explained as intra-allelic events. In general a trend to lose the variant repeat TGAGGG and homogenization to a pure array of TTAGGG was observed. The simplest explanation, consistent with the known function of BLM is that there was an increase in unequal sister chromatid exchange. However copying from a non-homologous telomere or extrachromosomal DNA of (TTAGGG) $_n$ repeats or deletion events cannot be ruled out.

In addition to the analysis presented here, previous analysis by Dr. Mendez-Bermudez in Dr. Royle's laboratory on the W-V cell line revealed decreased colony formation after transfection with the shBLM vector. Furthermore the W-V shBLM clone 6 showed an

increase in the mutation frequency in the minisatellite MS32 (~six fold) when compared with five different W-V control clones (Mendez-Bermudez *et al* paper submitted). Interestingly analysis of the pattern of tandem repeats in some of the MS32 mutant molecules showed the homogenization of the repeat array, similar to the mutations seen at the XpYp telomere (Mendez-Bermudez *et al*, paper submitted).

It has been proposed that the WRN and BLM helicases also have ssDNA-annealing activity, and it is thought that this activity might facilitate strand migration and ssDNA annealing during recombination or replication fork movement at the sites of DNA damage *in vivo* (Wu L *et al.* 2006), (Machwe A *et al.* 2005), (Cheok CF *et al.* 2005). WRN and BLM can also participate in resolving Holliday junction (HJ) structures and other structures formed during replication (e.g. G-quadruplex) (Bohr VA *et al.* 2008). The evidence presented here suggests that WRN participates in the generation of “complex mutations”, and because WRN is the only human RECQ helicase with nuclease activity it may be that this activity contributes to the generation of “complex mutations”. This could explain the absence of such mutations in the W-V cell line. Another explanation for the lack of “complex mutations” in the absence of WRN could be related to the formation of the complex between WRN and the RAD52 recombinase which is important for ALT+ cells (Bayton K *et al.* 2003). The increase in telomere mutation rate after downregulation of BLM could be related to its role in processing telomeric substrates or aberrant structures. This could cause replication fork stalling leading to an increase of unequal sister-chromatid exchange, as shown previously (Machwe A *et al.* 2005). BLM is an important member of protein complexes dedicated to replication fork progress in ALT cell lines (Temime-Smaali N *et al.* 2008, Bhattacharyya S *et al.* 2009). Reduced levels of BLM may disrupt complex formation leading to an ineffective processing of stalled replication forks and subsequent telomere mutation.

3.6 Future work

The data presented here support the importance of WRN and BLM helicases in the ALT pathway. In order to get a better understanding their roles in the ALT mechanism the following experiments could be done:

- Downregulation of WRN in an ALT+ cell line with endogenous expression of WRN. This experiment would help to confirm the participation of WRN in the generation of “complex mutations”.
- Expression of the WRN helicase in the W-V cell line. Even when it was tried in the present work it was not successful. The use of different vectors and delivery methods need to be attempted. In addition it would be interesting to transfect the W-V cell line with a nuclease deficient WRN mutant or a helicase deficient WRN mutant and observe the effects of each mutant on telomere and MS32 mutation frequency.

CHAPTER 4 – ROLE OF THE EXO1 EXONUCLEASE IN THE ALT PATHWAY

4.1 Background

Experiments in yeast strains with deletion of the gene *tlc1*, which encodes the RNA component of telomerase in yeast, have shown that some cells are able to survive by acquiring a recombination-based telomere maintenance mechanism. Two types of survivors can be observed, and both are dependent upon Rad52. In addition the type I survivors are dependent on Rad51 and they maintain their telomeres by amplifying the Y' subtelomeric elements. On the other hand the type II survivors are dependent on Rad50 and they maintain their telomeres by amplifying the TG₁₋₃ telomeric repeats. Experimental data have shown that *Exo1* is necessary for the develop of Type I and Type II survivors, which suggests that *Exo1* may be involved in the processing of the 5' end and the progression of the break induced replication (BIR) mechanism (Maringele L *et al.* 2004). Mutation of *Exo1* alone does not have any effect on telomere length, but in genetic backgrounds in which the telomeres are uncapped, *Exo1* is able to degrade the 5' end of the chromosome (Tran PT *et al.* 2004). The findings from experiments in yeast suggest two possible roles of *Exo1* at telomeres. First, when telomerase is functional *Exo1* seems to counteract telomere elongation, but, when either, the telomere is uncapped or telomerase is absent, *Exo1* seems to facilitate telomere maintenance by recombination between telomeric regions of two chromosomes. In this scenario *Exo1* has the opportunity to produce extensive ssDNA tails that activate a DNA damage checkpoint and promote recombination-driven lengthening of telomeres (Tran PT *et al.* 2004). A role of *Exo1* in the induction of checkpoints in response to telomere dysfunction has been revealed by studies in yeast. In telomere-dysfunctional strains the deletion of *Exo1* rescued yeast from the induction of the senescence (Hackett JA *et al.* 2003).

The function of EXO1 in telomere dysfunction in mammalian cells is not well understood. In the telomerase negative (mTerc^{-/-}) mouse model, the third generation (G3) show

reduction of cell proliferation and life span, and an increase in the levels of apoptosis and organ atrophy. Deletion of *Exo1*, in G3 mTerc^{-/-} mice resulted in improvement of the life span and cell proliferation, and there was an increase in the number of telomere-free ends and anaphase bridges. The effects observed after the deletion of *Exo1* seem to be related to the lack of generation of the 3' overhang and the subsequent activation of DNA damage response and cell cycle arrest, induced by ATR (Schaetzlein S *et al.* 2007).

The fact that some authors have proposed that the ALT mechanism is very similar to the BIR pathway in yeast, and that EXO1 seems to play an important role in progression of BIR makes EXO1 an attractive candidate to participate in the ALT pathway. To date it is not known if EXO1 is present in the APBs in ALT cells, but there is evidence that some of the proteins found in the APBs are able to interact with EXO1. Among these are the WRN and BLM helicases (Henson JD *et al.* 2002). Some studies have shown that the COOH-terminal region of WRN is able to interact functional and physically with EXO1, modulating in a positive way its cleavage activity, and this interaction seems to be independent of the helicase activity of WRN (Sharma S *et al.* 2003). Exonuclease assays have demonstrated that BLM stimulates dsDNA resection by hEXO1 through a physical interaction between the proteins. No difference was observed in the stimulation of both isoforms of EXO1, suggesting that the last 43 amino acids in the C-terminal of *Exo1* are not essential for its interaction with BLM (Nimonkar AV *et al.* 2008). Experiments *in vitro* using different substrates containing telomeric repeats have demonstrated that the exonuclease activity of EXO1 can be stimulated to process the C strand of a telomeric substrate compared to substrates with no telomeric repeats (Vallur AC *et al.* 2010). So far the role of EXO1 has not been investigated in humans in ALT cells.

Previous experiments carried out in Dr. Royle's laboratory suggested higher levels of expression of EXO1 in ALT+ cell lines compared to telomerase+ cell lines (Novo C, *et al.* unpublished data). This finding, plus the evidence reported in the literature stimulated my investigation of the possible role of EXO1 in the ALT pathway.

4.2 Aim

In this chapter I have attempted to explore the role of EXO1 in the ALT mechanism in human cell lines.

4.3 Analysis of *EXO1* sequence in telomerase positive and ALT positive cell lines

The EXO1 gene is highly polymorphic in humans with more than 600 SNPs reported (www.ncbi.nlm.nih.gov/sites/entrez/EXO1). Studies in patients with Lynch syndrome (also known as HNPCC) have found that some base substitutions can reduce the ability of EXO1 to interact with components of the MMR machinery (Sun X *et al.* 2002), (Wu Y *et al.* 2001), in addition there is evidence that certain amino acids are important for its degradation via the proteasome (El-Shemerly M *et al.* 2007). In order to investigate whether ALT+ cell lines contain a combination of coding SNPs that could interfere with the enzymatic function of EXO1 the sequence of the 13 coding exons and the promoter region was determined in eight cell lines. WI38 is a normal cell line derived from lung fibroblasts, the JFCF/6-T.1J/11C cell line express telomerase and the rest of the cell lines use the ALT mechanism for telomere maintenance. The primers used were described previously (Jagmohan-Changur S *et al.* 2003).

The analysis revealed 19 SNPs in the gene, 18 known and one unreported SNP. Seven of the SNPs were found in the coding sequence, two in the promoter region and ten in introns, but not at exon-intron boundaries (Table 4.1). None of the SNPs observed lies in the nuclease domains, and none of them have been associated with a reduction in the exonuclease activity or interaction with the members of the MMR system.

The cell lines JFCF/6-T.1J/11C and JFCF/6-T.1J/11E derive from a common ancestor, as do the cell lines WI38 and WI38VA13/2RA. It is important to notice that the genotypes are conserved regardless of the telomere maintenance mechanism used. This finding suggests that there is no relationship between EXO1 mutations and ALT activity. The cell lines

SAOS, U2OS, and IICF were homozygous for all the SNPs identified. The deletion of one allele of the *EXO1* gene could explain this observation. However, no additional experiments were conducted to determine whether the cell line showed loss of heterozygosity for *EXO1*.

From the SNPs present in the coding sequence one is located in the interaction domain with MSH3, one in the interaction domain with MLH1 and three in the interaction domain with MSH2 (figure 4.1). The unreported SNP lies in intron four, nucleotide position 4286 (accession number NC_000001.10) and within this small sample of cell lines the frequencies for the alleles G and A are 0.75, and 0.25 respectively. A larger survey of allele frequencies at this SNP was not undertaken, because there is no evidence that this polymorphism would affect *EXO1* function.

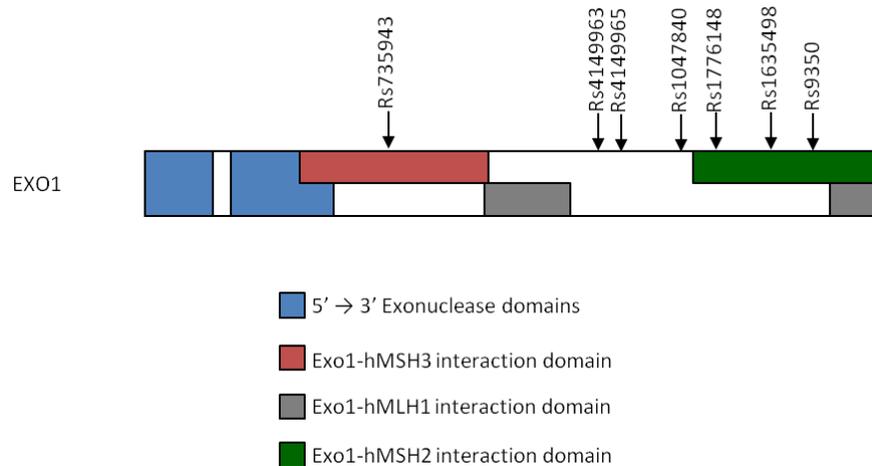


Figure 4.1. Known SNPs in the coding sequence of *EXO1*. *EXO1* contains two exonuclease domains (in blue) and interaction domains for three members of the MMR system hMSH3 (in red), hMLH1 (in gray) and hMSH2 (in green). SNP Rs735943 lies on exon 10 in the hMSH3 interaction domain, SNPs Rs4149963 and Rs4149965 lie in exon 11, SNPs Rs1047840 and Rs1776148 lie in exon 12, but only Rs1776148 is in the hMSH2 interaction domain. Two further SNPs, Rs1635498 in exon 13 and Rs9350 in exon 14 also lie in the hMSH2 interaction domain.

Table 4.1 SNPs found in *EXO1* in eight different cell lines. WI38 is a normal cell line derived from lung fibroblasts, the JFCF/6-T.1J/11C cell line express telomerase and the rest of the cell lines use the ALT mechanism for telomere maintenance. In the first row is the SNP ID, the region where the SNP lies, the alternative alleles with their frequencies reported for European population and the amino acid change when the SNP lies in coding sequence. The column Geno shows the genotype of the cell line and the column aa the amino acid if the SNP is coding. SNP shown in bold are located in coding sequence. The SNPs lying on exon 1 and exon 15 belong to the 5' and 3' UTR respectively. *SNP previously unreported, nucleotide position according to the gene bank accession number NC_000001.10.

	Rs1635517 Promoter G/A 0.44/0.56		Rs776178 Promoter A/G 0.4/0.6		Rs1776177 Exon 1 C/T 0.5/0.5		Rs4149867 Intron 4 C/T 0.87/0.13		Intron 4 Nucleotide 4286*		Rs4149896 Intron 6 A/G 0.87/0.13		Rs2526699 Intron 7 T/C 0.43/0.56		Rs735943 Exon 10 A/G 0.33/0.67 aa H/R		Rs4149963 Exon 11 C/T 0.95/0.05 aa T/M	
	Geno	Aa	Geno	aa	Geno	aa	Geno	aa	Geno	Aa	Geno	aa	Geno	aa	Geno	aa	Geno	aa
WI38	G/A	-	G/G	-	C/C	-	C/T	-	A/G	-	A/G	-	T/T	-	A/G	H/R	C/T	T/M
JFCF/6-T.1J/11C	A/A	-	G/G	-	C/T	-	C/C	-	G/G	-	A/G	-	C/C	-	G/G	R/R	C/T	T/M
JFCF/6-T.1J/11E	A/A	-	G/G	-	C/T	-	C/C	-	G/G	-	A/G	-	C/C	-	G/G	R/R	C/T	T/M
WI38V13/2RA	A/A	-	G/G	-	C/C	-	C/T	-	A/G	-	A/G	-	T/T	-	A/G	H/R	C/T	T/M
IIICF/c	G/G	-	A/A	-	C/C	-	C/C	-	A/A	-	A/A	-	T/T	-	A/A	H/H	C/C	T/T
U2OS	A/A	-	G/G	-	T/T	-	C/C	-	G/G	-	A/A	-	C/C	-	G/G	R/R	C/C	T/T
SAOS	A/A	-	G/G	-	T/T	-	C/C	-	G/G	-	A/A	-	C/C	-	G/G	R/R	C/C	T/T
SUSM-1	G/G	-	G/G	-	C/T	-	C/C	-	G/G	-	A/A	-	C/C	-	A/G	H/R	C/C	T/T

	Rs4149965 Exon 11 G/A, 0.9/0.1 aa V/M		Rs1047840 Exon 12 G/A 0.66/0.34 aa E/K		Rs1776148 Exon 12 A/G 0.32/0.68 aa E/G		Rs1418760 Intron 12 A/G 0.24/0.76		Rs1635498 Exon 13 C/T 0.18/0.82 aa R/C		Rs9350 Exon 14 C/T 0.67/0.33 aa P/L		Rs4150020 Intron 14 C/G 0.69/0.31		Rs4150021 Exon 15 T/- 0.7/0.3		Rs851797 Exon 15 A/G 0.37/0.63		Rs4150024 Exon 15 G/T 0.99/0.01	
	Geno	Aa	Geno	aa	Geno	aa	Geno	aa	Geno	aa	Geno	aa	Geno	aa	Geno	Aa	Geno	aa	Geno	aa
WI38	G/G	V/V	G/A	E/K	G/G	G/G	G/G	-	T/T	C/C	T/T	L/L	G/G	-	-/-	-	A/A	-	G/G	-
JFCF/6-T.1J/11C	G/A	V/M	G/A	E/K	A/G	E/G	A/G	-	T/T	C/C	C/T	P/L	C/G	-	T/-	-	A/A	-	G/G	-
JFCF/6-T.1J/11E	G/A	V/M	G/A	E/K	A/G	E/G	A/G	-	T/T	C/C	C/T	P/L	C/G	-	T/-	-	A/A	-	G/G	-
WI38V13/2RA	G/G	V/V	G/A	E/K	G/G	G/G	G/G	-	T/T	C/C	T/T	L/L	G/G	-	-/-	-	A/A	-	G/G	-
IIICF/c	G/G	V/V	G/G	E/E	G/G	G/G	G/G	-	T/T	C/C	C/C	P/P	C/C	-	T/T	-	G/G	-	T/T	-
U2OS	G/G	V/V	G/G	E/E	G/G	G/G	G/G	-	T/T	C/C	C/C	P/P	C/C	-	T/T	-	G/G	-	G/G	-
SAOS	A/A	M/M	G/G	E/E	G/G	G/G	G/G	-	T/T	C/C	T/T	L/L	G/G	-	-/-	-	A/A	-	G/G	-
SUSM-1	G/G	V/V	A/A	K/K	G/G	G/G	G/G	-	T/T	C/C	T/T	L/L	G/G	-	-/-	-	A/A	-	G/G	-

4.4 Analysis of EXO1 expression by Western blot in telomerase+ and ALT cell lines

Previous analysis by real time PCR (qPCR) had shown higher expression of *EXO1* in ALT cell lines when compared to telomerase positive or normal cell lines (telomerase-/ALT-) (Novo C, unpublished data), suggesting that EXO1 could be playing a role in the ALT pathway. In order to use a different approach to analyse the expression of EXO1 Western blot analysis was carried out with a commercial goat-polyclonal antibody from Santa Cruz (sc-19941) raised against the N-terminus of the human protein that should recognize both isoforms of EXO1 (Bolderson E *et al.* 2009), and a secondary antibody (donkey anti-goat IgG-HRP sc-2020) that was recommended by the manufacturer. As a control an antibody against the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used. Initially 80µg of total protein extract from three different ALT positive cell lines were used. According to the manufacturer the EXO1 protein should be ~100kDa but antibody detection showed multiple bands ~100kDa and a very strong band ~50kDa (figure 4.2) and it was not possible to determine which band corresponded to EXO1. Different blocking buffers, incubation times, and increasing the amount of protein loaded (100µg of total protein) were tried but the specificity of the antibody did not improve (data not shown). As a positive control mouse testis was used because it is reported to have high levels of Exo1 expression and the antibody is supposed to react to it. The mouse Exo1 protein contains 837 aminoacids, and has a similar molecular weight to the human protein; 100µg of total protein and 100µg of nuclear extract were tested but without improvement. In all the cases the control Western for GAPDH gave a clear, strong signal (data not shown). The lack of Exo1 antibody specificity prevented the use of Western blot as a reliable method to evaluate *EXO1* expression.

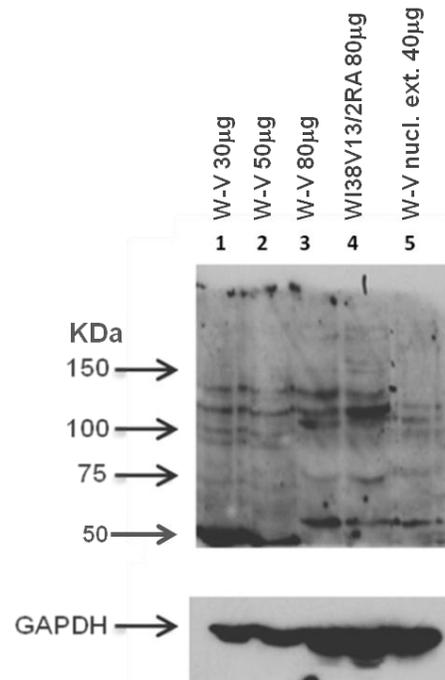


Figure 4.2 EXO1 Western blot. Lane 1, 2 and 3 cell extracts from the W-V cell line 30 µg 50 µg and 80 µg respectively, lane 4 WI38V132RA cell line extract 80 µg, lane 5 W-V nuclear extract 40 µg. Bottom panel shows the GAPDH Western blot as a loading control.

4.5 Expression analysis of EXO1 by real time PCR in telomerase+ and ALT cell lines

Failure to obtain satisfactory results by Western blot analysis lead me to repeat the *EXO1* expression analysis by qPCR, using GAPDH and β -Actin as reference genes. Primers for each gene spanning exon/intron boundaries were designed previously (Novo C unpublished data). A 10 fold cDNA dilution series was used to calculate the efficiencies for each amplicon using the formula $(10^{-1/\text{slope}}-1)$, and then the efficiencies of the different amplicons were compared between them in order to determine whether the primer pairs had similar efficiencies (Bustin SA *et al.* 2009). The efficiencies obtained for the three amplicons were between 0.9 and 1 which were considered acceptable, and the comparison between efficiencies showed slope values <0.1 confirming that all primer pairs had similar efficiencies, see appendix 3.

Relative expression of *EXO1* was analysed in the telomerase+ cell lines HT1080 and JFCF/6-T.1J/11C and in the ALT cell lines JFCF/6-T.1J/11E, WI38VA13/2RA, U2OS, SAOS, and SUSMI. The lymphoblastoid cell line from a normal donor GMO3798 was used as calibrator and the *EXO1* Ct values were normalised against the reference genes GAPDH or β -Actin, as previously done (Novo C unpublished data). The analysis did not show a higher expression of *EXO1* in the ALT positive cell lines (Figure 4.3). The JFCF/6-T.1J/11C telomerase+ cell line showed slightly higher expression than the ALT cells, except for WI38V13/2RA. It is worth mentioning that the WI38V13/2RA cell line contains an extra copy of chromosome 1, where *EXO1* lies (N J Royle unpublished data) and this could contribute to the higher *EXO1* expression observed in this cell line.

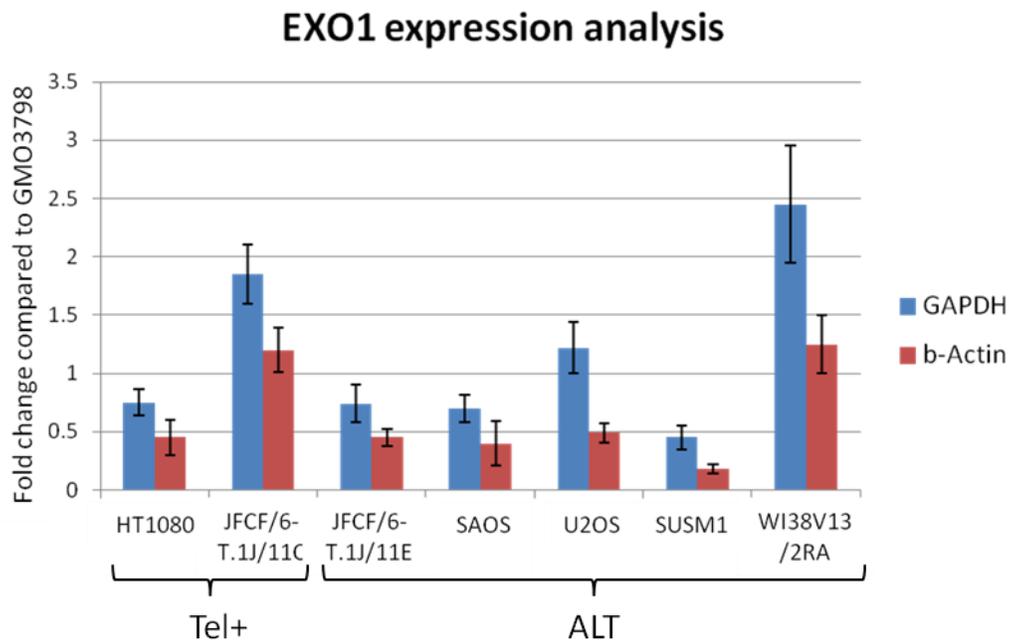


Figure 4.3 Relative expression of *EXO1* in the telomerase+ cell lines HT1080 and JFCF/6-T.1J/11C and the ALT cell lines JFCF/6-T.1J/11E, WI38VA13/2RA, U2OS, SAOS, and SUSMI. The lymphoblastoid cell line GMO3798 was used as calibrator, GAPDH and β -Actin were used as reference genes. All the reaction were done in triplicate

4.6 Downregulation of EXO1 in ALT positive cell lines

In spite of the lack of reproducibility of the qPCR results, the probability of EXO1 playing a role in the ALT pathway is supported by the evidence that EXO1 is required for the establishment of type II survivors in yeast studies (Maringele L *et al.* 2004). To determine whether downregulation of *EXO1* affects ALT cell lines the SUSM1 and WI38VA13/2RA cell lines were transfected with the pSUPERIOR vector expressing a shRNA against *EXO1* (shEXO1) or an empty vector as a control. Two different versions of shEXO1 were generated, shEXO1A and shEXO1B. Each was designed to target both *EXO1* transcripts. The cell lines were transfected using different methods to try to achieve good transfection efficiencies. The WI38VA13/2RA cells were transfected using the FuGENE reagent (Promega) and the SUSM1 cells were transfected by electroporation (Gene Pulser Xcell™ electroporation system Bio-Rad).

4.6.1 Colony forming assay in ALT cells transfected with shRNA against *EXO1*

For the SUSM1 cell line three million cells were electroporated with each plasmid and the electroporated cells were divided among three 10cm Petri dishes each containing 15ml of media. For the WI38V13/2RA one million cells were plated in 10cm petri dishes in 15ml of medium 24 hrs before transfection with the FuGENE reagent (Promega). After transfection cells were cultured in media with G418 (500µg/ml) to select for cells that contained the pSuperior vector and after 14 days the plates were fixed and stained with 1% crystal violet and colonies counted. For the WI38VA13/2RA cells transfected with the shControl, shEXO1A or shEXO1B plasmids the number of colonies counted were 33, 51 and 3 respectively. For the SUSM1 cells transfected with the shControl, shEXO1A or shEXO1B plasmids were 8, 13, and 5 respectively (Figure 4.4). The A version of the shEXO1 seemed to have no effect on colony forming efficiency (actually there was an increase in the number of colonies), but the B version caused a substantial reduction in the colony forming efficiency in the WI38VA13/2RA cells and a modest reduction in the SUSM1 cells

lines. Whether the differences between version A and B were due to better efficiency of *EXO1* downregulation by version B or due to an off-target effect was not determined.

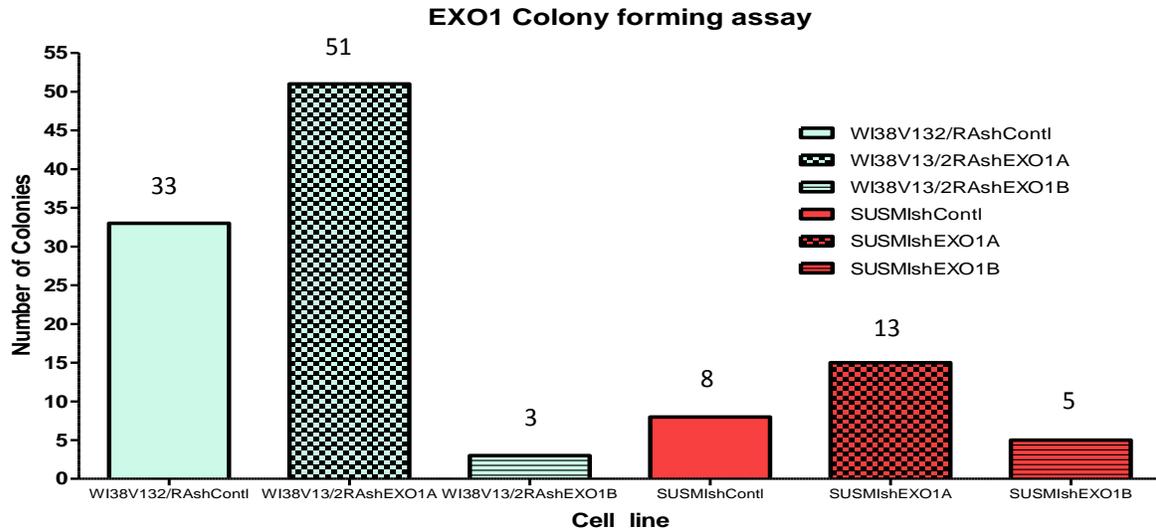


Figure 4.4 Colony forming assay in ALT+ cell lines transfected with shEXO1 plasmids. The colonies were counted 14 days after transfection of the cells with a shControl, shEXO1A or shEXO1B plasmid.

4.6.2 Expansion of clones after transfection with a shControl, shEXO1A or shEXO1B.

Besides the analysis of colony formation efficiency, some colonies were isolated and expanded as described in the methods section. Four control clones and one shEXO1A clone were recovered for WI38V13/2RA, but no colonies were recovered from the shEXO1B transfected cells. Whereas five control clones, eight shEXO1A clones and four shEXO1B clones were recovered for SUSM1. Population doublings were calculated on every passage and growth curves were generated (Figure 4.5). An important growth delay of the WI38V13/2RA shEXO1A clone was observed when compared to controls, it took more than 70 days to reach PD15, whereas control clones reached approximately the same point at day 30 (Figure 4.5A). On the other hand the growth curve for SUSM1 clones showed that control clones reached PD20 after approximately 30 days, but the growth rates of the treated clones were very heterogeneous, for example shEXO1B reached the

same PD basically at the same time as controls, whereas shEXO A9, shEXO A11, shEXO A13 and shEXO B2 clones got to that point at around day 45. The slowest growing SUSM1 clones were shEXO A5, shEXO A3, shEXO A6, shEXO B3 and shEXO B4, taking more than 60 days to reach PD20. To determine whether the observed growth delay could be related to reduced *EXO1* expression, qPCR was carried out.

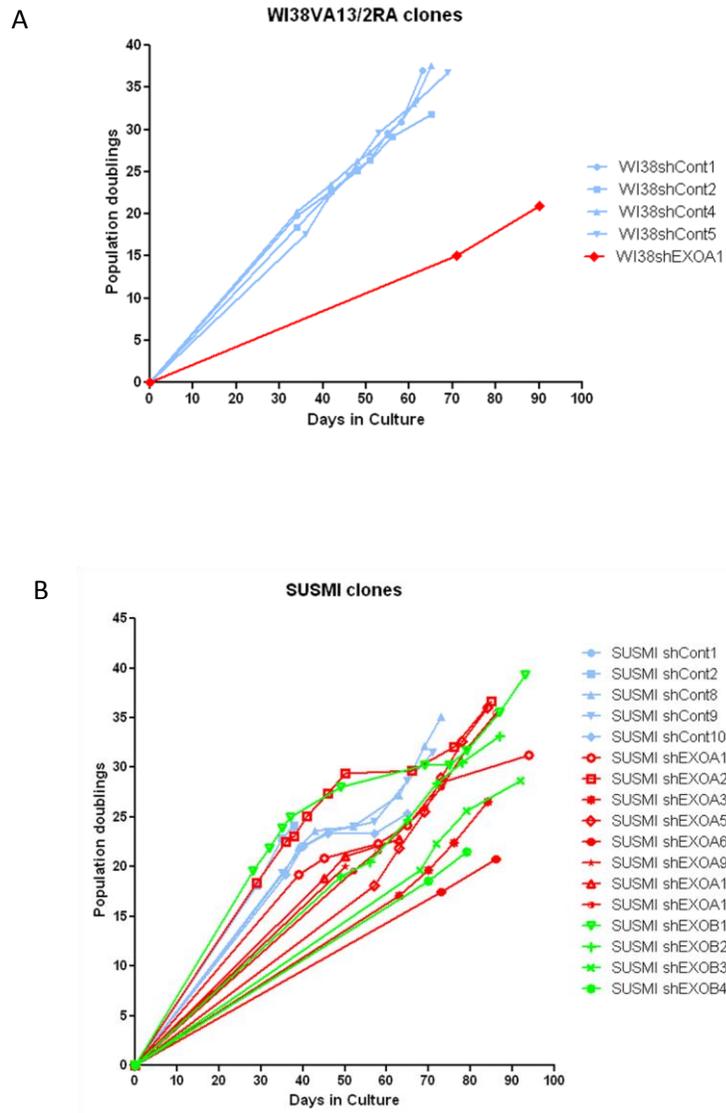


Figure 4.5 Growth curves of the clones expanded after transfection of the ALT+ cell lines WI38V13/2RA (A) and SUSM1 (B) with a shControl (blue), shEXO1A (red) or shEXO1B (green) plasmid.

As described in the methods chapter (section 2.11) the geNorm reference gene selection kit (Primer design) and the geNorm applet included in the qbase^{plus} software (Biogazelle) were used to select the reference genes with the most stable expression for each cell line. The principle for reference gene selection starts from the idea that if the expression ratio between two genes remains constant in a given sample, regardless of experimental conditions, they are considered as genes with stable expression. The geNorm applet determines which gene or genes have the most stable expression within a sample, by making a comparison of each gene against each one of the genes included in the analysis, and all the possible combinations are evaluated. The result is the identification of the genes that show constant expression ratios. These genes are considered to have the most stable expression and so they can be used for normalization in the sample (Vandesompele *J et al.* 2002). The genes included in the geNorm analysis were obtained from the geNorm reference gene selection kit (Primer design) (see Table 2.7). In order to identify the genes with the most stable expression in this study the genes included in the geNorm reference gene selection kit were tested in eight different cDNAs obtained from eight different cell pellets from each cell line. According to the geNorm reference gene selection kit and the qBASE software (Biogazelle) the reference genes needed for WI38V13/2RA were β -Actin, GAPDH and the succinate dehydrogenase complex (SDHA). For SUSMI the selected genes were ubiquitin C (UBC) and ATP synthase, (ATP5B) (Appendix 3).

After selection of the adequate reference genes for each cell line, RNA was extracted from each clone at ~PD20 and after cDNA synthesis, using the VersoTM cDNA Kit (Thermo), qPCR was carried out to check expression of *EXO1*. All qPCR reactions were set up in triplicate. In addition to comparing the *EXO1* expression levels between transfected clones, cDNA from the progenitor cell lines was included in the qPCR analysis.

The qPCR results in WI38V13/2RA showed that at ~PD20 the shEXO1 clone did not show reduced expression of *EXO1* compared to the progenitor cell line and control clones (Figure 4.6). The analysis was repeated on the same cDNA showing the same pattern of expression. This result was unexpected given the slow growth rate of the WI38V13/2RA

shEXO1 clone but it is possible that EXO1 expression was reduced at an earlier stage in the expansion of the clone. In fact the WI38V13/2RA shEXO1 clone reached PD15 at day 71 in culture, this corresponds to a growth rate of 0.21PD/day, whereas the control clones had an average growth rate of 0.52PD/day which is more than twice faster. However over the next 19 days the shEXO1 clone reached PD 20.9, and so the growth rate increased to 0.31PD/day. It is possible that the increased growth rate was related to recovery of *EXO1* expression, unfortunately this is difficult to prove because before PD15 there were not enough cells to extract RNA, DNA and keep the cells in culture.

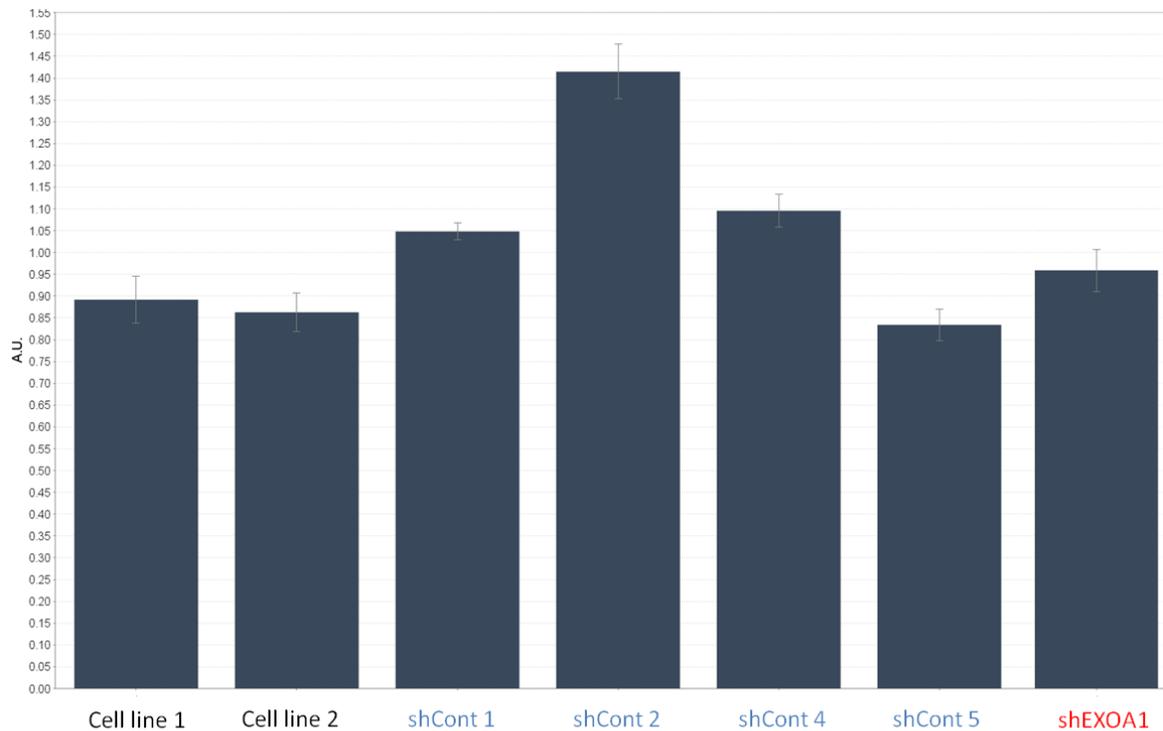


Figure 4.6 *EXO1* qPCR expression analysis between the progenitor WI38V13/2RA cell line (Cell line 1 and Cell line 2), control clones and shEXO1 clone. Data was normalised using β -Actin, GAPDH and SDHA as reference genes using the qbase^{plus} software. All the reactions were set up in triplicate. A.U. arbitrary units.

The qPCR gene expression analysis results of the SUSM1 clones are shown in Figure 4.7. The results of the SUSM1 shEXO (A1, A2, A5, A11, A13, B1 and B3) clones which reached ~PD20 first plus the control clones and two different cell pellets of the progenitor cell line, are shown in Figure 4.7A. In this figure the variability of *EXO1* expression among control clones can be observed clearly. For example, SUSM1 control clones one and two have similar expression levels compared to the progenitor cell line; whereas control clones eight and nine show considerably lower levels of expression. This variation in *EXO1* expression levels among control clones makes interpretation of the *EXO1* expression data from clones transfected with shEXOA or B difficult. Consequently, analysis of *EXO1* expression in the SUSM1 shEXOA and shEXOB clones did not show clear evidence of downregulation of *EXO1* expression when compared against the control clones, mainly because of the low expression of control clones eight and nine. Figure 4.7.B shows the qPCR analysis results of the slow growing SUSM1 shEXO clones (A3, A6, B2 and B4) plus control clones and the progenitor cell line. In addition this figure shows a replica of the analysis done on the same cDNA from the SUSM1 shEXO clones A1, A3, A5, A11, B1 and B3 in order to confirm the results shown in Figure 4.7A. On this occasion the SUSM1 control clones one and two behaved in a very different way showing very low levels of expression compared to the progenitor cell lines, and control clone eight was the one with similar levels of expression to those observed for the progenitor cell lines. The lack of reproducibility of the qPCR assay for *EXO1*, even with the inclusion of at least two control genes, meant that it was not possible to determine whether either short hairpin construct (shEXOA or B) was effectively reducing *EXO1* expression. (For technical replicates see appendix 4).

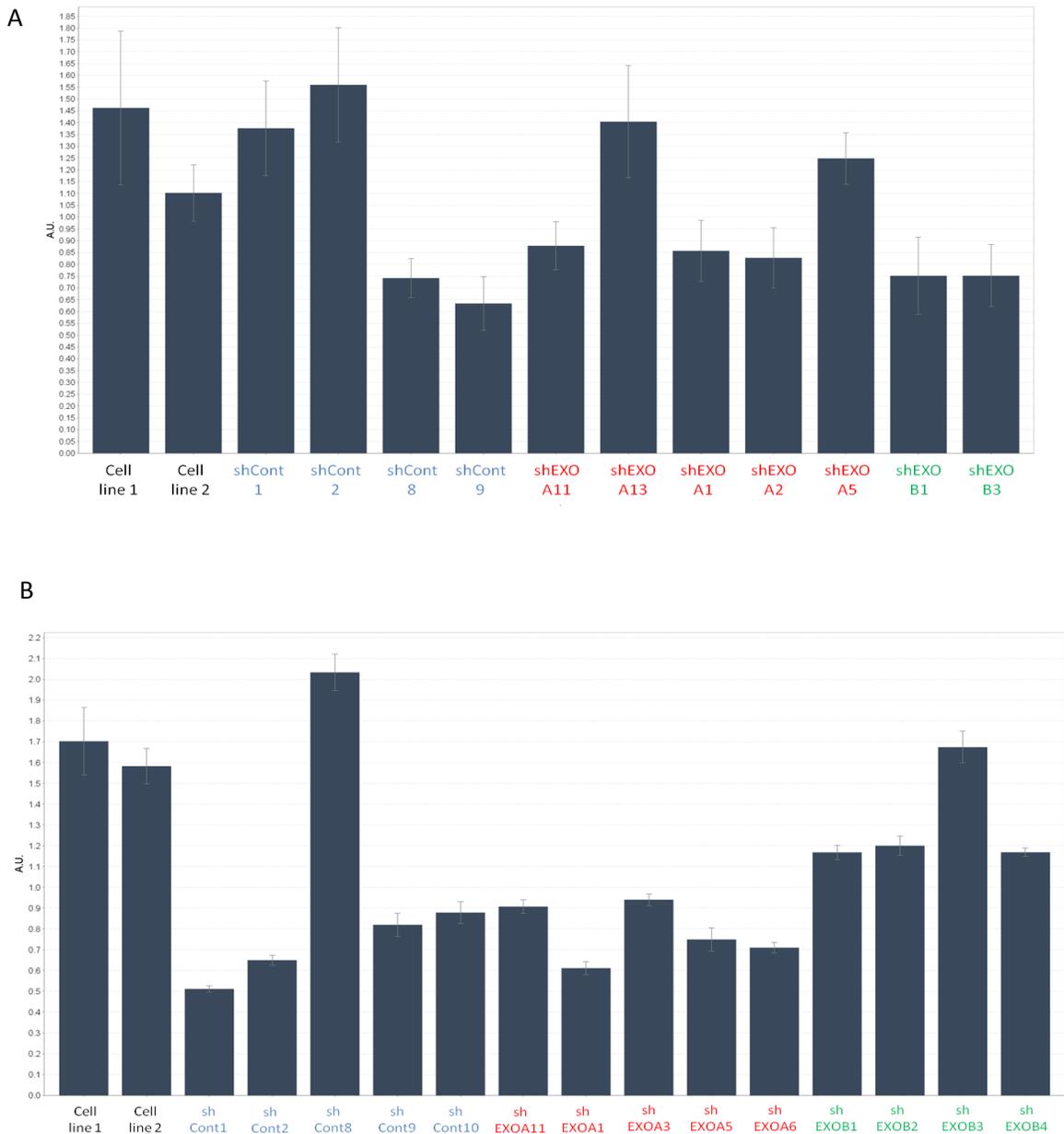


Figure 4.7 Comparison of the *EXO1* expression analysis between the progenitor SUSM1 cell line, control clones and shEXO A and shEXO B clones. (A) Shows the results of the qPCR carried out in the SUSM1 clones that reached PD20 first (A1, A2, A5, A11, A13, B1 and B3) plus control clones and the progenitor cell line. (B) Shows the results of the qPCR carried out in the slow growing SUSM1 clones (A3, A6, B2 and B4), plus a replica of on the same cDNA of the clones A1, A3, A5, A11, B1 and B3, control clones and the progenitor cell line were also included in the analysis. All qPCR reactions were set up in triplicate. Data was normalised using ATP5 and UBC as reference genes using the qbase^{plus} software. A.U. arbitrary units. For experimental replica see appendix 4.

However, it remains possible that the growth delay observed in SUSM1 clones (particularly SUSM1 shEXO6A or shEXO6B4, Figure 4.5B) at early stages of clonal expansion is a consequence of EXO1 downregulation which was then lost. Evidence supporting this is that in both cell lines control clones grew at a very similar rate. The lack of strong evidence that EXO1 expression had been downregulated in clones from two ALT+ cell lines led to the decision to discontinue this line of investigation and so further experiments to investigate telomere function were not initiated.

4.7 Discussion

EXO1 has been extensively studied as an interacting partner of the MMR system and as part of the DSB repair machinery. Emerging evidence suggests that the same machinery is involved in the telomerase-independent TMM. It is known that some aminoacids are important for the correct exonuclease function of EXO1 and some, via phosphorylation, regulate its participation in DSB repair (Bolderson E *et al.* 2009), (Morin I *et al.* 2008). The sequence of the promoter and coding sequence of *EXO1* in six ALT+ cell lines, one telomerase+ and one normal cell line (ALT-/telomerase-) revealed seven SNPs in the coding sequence but none of them were located in the nuclease domains. There is no strong evidence to suggest that these SNPs could interfere with the EXO1 activity or interactions with other proteins partners and so there is no indication of a relationship between the TMM and known SNPs in the *EXO1* gene.

Western blot analysis was attempted to corroborate the previous data suggesting there was a higher *EXO1* expression in ALT cells compared to telomerase positive cells (Novo C, unpublished data). However the lack of antibody specificity prevented the use of this technique to evaluate the expression of EXO1. Some authors have reported that EXO1 expression is very low in some cell lines, making it difficult to detect the protein by Western blot analysis (El-Shemerly M *et al.* 2005). Attempts to reproduce the data by qPCR did not show any evidence of higher expression of *EXO1* in ALT cells. It is worth

mentioning that the expression was measured in asynchronously growing cultures, but as ALT activity is thought to occur in late *S/G2*, it would be interesting to measure *EXO1* expression in synchronized cultures to confirm or discard the suggest that there are differences in expression between cells using ALT, telomerase or no TMM.

In telomerase negative budding yeast strains the emergence of type I and type II survivors is delayed in the absence of Exo1. This suggests that Exo1 has an important but not essential role in the generation of survivors (Maringele L *et al.* 2004). To investigate the consequences of *EXO1* depletion in ALT cell survival, the WI38V13/2RA and SUSMI ALT+ cell lines were transfected with one of two shRNAs designed against *EXO1*. The ability to form colonies was measured, and only one of the shRNA (B) resulted in a decrease in the number of colonies formed, particularly in the WI38V13/2RA cells, raising the possibility that *EXO1* may play an important role in ALT+ cells.

To investigate whether the reduction in colony formation was related to *EXO1* depletion and telomere dysfunction attempts were made to generate clones with stable expression of an shEXO1 targeting *EXO1* expression. The recovered clones where expanded and *EXO1* expression measured by qPCR at ~PD20. None of the analysed clones from either cell line showed a clear pattern of downregulation of *EXO1* expression at the point of analysis. The growth curves of the clones showed that control clones grew at similar rates (0.52 PD/day and 0.57PD/day for WI38VA13/2RA and SUSMI respectively). The only WI38V13/2RA shEXO1 clone recovered had a growth rate of 0.21PD/day at PD15, but after that the growth rate increased to 0.31PD/day. Among the SUSM1 clones, some had similar growth rates compared to control clones (A1, A2, A9, A11, A13, B1 and B2). However, other SUSM1 shEXO clones clearly showed reduced growth rates (A3, A5, A6, B3 and B4), so it is possible that in these clones cell growth could be affected by *EXO1* downregulation at early stages of clonal expansion. Whether telomere instability contributes to this reduced growth rate remains to be answered. In addition evidence from yeast suggests that Exo1 could have a redundant role in 5' processing of the C-strand with at least one other nuclease. According to the model proposed by Bonetti *et al.*, the nuclease Dna2 could

generate long 3' tails (Bonetti D *et al.* 2009). A different approach will be needed to study the role of EXO1 on telomere maintenance in ALT+ cell lines.

4.8 Future work

Some of the experiments that could help to elucidate EXO1's role in the ALT pathway are:

- Association of EXO1 with telomeres and APBs using a fluorescent tagged EXO1. So far it is not known if EXO1 is present in APBs, or if it is more frequently associated with telomeres in ALT+ cells than in telomerase+ or normal (ALT-/telomerase-) cells. It would be important to investigate whether EXO1 is associated with telomeres and/or APBs and at which phase of the cell cycle. To overcome the lack of a reliable commercial antibody against EXO1, cells could be transfected with a vector expressing a fluorescent tagged EXO1. Some authors have successfully used this strategy (YFP-EXO1) to study the import of EXO1 into the nucleus (Knudsen NO *et al.* 2007), (Andersen SD *et al.* 2012).
- Expression analysis by qPCR in synchronized cultures. ALT activity is thought to occur in late S/G2. The comparison of the expression levels between ALT+, telomerase+ and normal (ALT-/telomerase-) cells in synchronized cultures would provide data about the phase where EXO1 is preferentially expressed according to the TMM used. However, before doing this a robust qPCR assay to measure EXO1 expression must be established. The use of different primers for EXO1 could potentially help to improve the results of the qPCR assay.
- Transient downregulation of EXO1 using siRNA. This experiment could be useful to look for rapid effects on ALT+ cell lines upon transient downregulation of EXO1 such as telomere fusions or telomere free ends. The use of telomerase+ or normal (ALT-/telomerase-) cell lines as controls would be necessary to establish that an observed effect is specific for cells using the ALT mechanism.
- Stable transfection using viral vectors. In this study it was not possible to generate

clones from the ALT+ cell lines with stable downregulation of EXO1, this could be caused by a poor transfection efficiency or a inefficient expression of the shRNAs. In order to generate clones with sustainable downregulation different transfection approaches should be considered. Viral vectors offer high transfection efficiencies and should be considered.

CHAPTER 5 – ROLE OF CtIP IN THE ALT PATHWAY

5.1 Background

Little has been done to investigate the possible role of CtIP in telomere maintenance or in the ALT pathway. CtIP's physical and functional interactions with the MRN complex and EXO1 make it an attractive candidate for participating in a HR based TMM (Huertas P *et al.* 2008). Some data supporting this idea came from studies in *S. cerevisiae* where the CtIP homolog *Sae2* was found to participate in the formation of the 3' overhang at telomeres which is crucial for maintaining telomere stability. An inducible short telomere assay was developed by Bonetti *et al.* to detect *de novo* telomere addition. In this system a yeast chromosome is modified so that after digestion with a specific endonuclease a short telomere is generated at one chromosome end, allowing monitoring of telomere extension. Results of the study proved that in telomerase positive yeast deletion of *Sae2* impaired telomere elongation but it was not completely abolished. The study also monitored degradation of the 5' end on the C-strand and degradation of the C strand was clearly reduced after deletion of *Sae2* in telomerase positive and negative yeast strains. This suggested that *Sae2* is required for ss 3' overhang production, similar to what happens in DSB repair. Interestingly telomere extension defects were rescued when cells overexpressed *Exo1* (Bonetti D *et al.* 2009). Previous investigations have showed that *Sae2* phosphorylation on serine 267 is needed for correct DSB processing. Since substitution of serine at position 267 for a non-phosphorylatable alanine residue caused reduction in telomere elongation and C-strand processing (Huertas P *et al.* 2008). This highlighted the importance of serine 267 for *Sae2* function. In order to investigate the genes involved in residual C strand processing double mutants combining *Sae2*, *Sgs1* or *Exo1* were generated. The *Sae2 Sgs1* double mutants showed the greatest defect. This suggests the existence of two pathways for 5' processing. In one of them, *Sae2* and the MRX complex initiate the process and then *Exo1* is recruited (Figure 5.1). The other one is driven by

Sgs1, with the participation of Exo1. Further analysis proposed that the exonuclease/helicase Dna2 can compensate the absence of Exo1. (Figure 5.1) (Bonetti D *et al.* 2009).

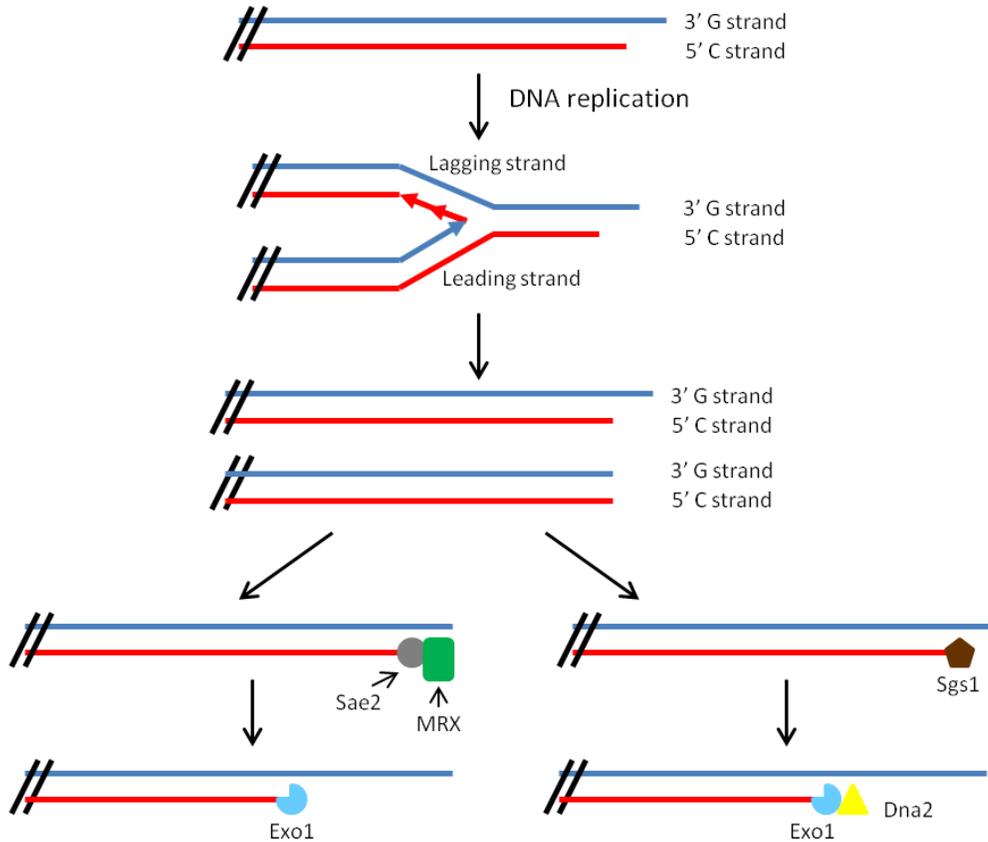


Figure 5.1 Model proposed for C-strand processing at telomeres in yeast after DNA replication. One pathway (left) is thought to be initiated by the coordinated action of Sae2 and the MRX complex with further processing of the 5' end by Exo1. The second pathway (right) is started by the helicase activity of Sgs1 and continued by Exo1 and/or Dna2. Taken from (Bonetti D *et al.* 2009).

Recent work by Kaidi A *et al.*, has demonstrated that CtIP is a substrate of the histone deacetylase (HDAC) SIRT6. The authors observed that downregulation on the expression of SIRT6 impaired CtIP deacetylation. The final consequence was a reduction of ssDNA

formation at DSBs, indicated by absence of RPA foci at the sites of the breaks. However, this effect was not observed when SIRT1 expression was downregulated or when cells were treated with inhibitors against other HDACs. These findings suggest a model where deacetylation of CtIP is required to generate 3' ssDNA tails for repair of DSB via HR (Kaidi A *et al.* 2010). SIRT6 has been shown to be important for telomere stability as in a mouse model null for SIRT6 an aging related phenotype was observed (Mostoslavsky R *et al.* 2006). In addition SIRT6 has been found to be associated with telomeric chromatin and it is responsible for deacetylation of at least two lysine residues on histone three (H3), H3K9 and H3K56 (Tennen RI *et al.* 2011). Depletion of SIRT6 in primary human fibroblast resulted in hyperacetylation on H3K9 and H3K56 at S phase, telomeric chromatin loss, accumulation of Telomere Induced Foci (TIFs) and end to end fusions. All these events triggered senescence. Furthermore site-specific histone deacetylation at telomeres is required for stable association with WRN (Michishita E *et al.* 2008), (Tennen RI *et al.* 2011). The interaction of SIRT6 with CtIP offers another link for a possible role of CtIP in the ALT pathway.

5.2 Aim

The Aim of this chapter was to investigate the role of CtIP in the ALT pathway.

5.3 Colocalization of CtIP with telomeres

To investigate if CtIP associates with telomeres in ALT or telomerase+ cell lines immunofluorescence was carried out in asynchronous cultures of the telomerase+ cell lines HT1080 and HeLa and the ALT cell lines WI38V13/2RA, SUSM1, W-V and SAOS. The probe was a polyclonal antibody against CtIP and a telomeric PNA probe labelled with Cyt 3. The PNA probe can hybridize to chromosomal and extrachromosomal telomeric DNA. All the cells analysed by confocal microscopy were positive for multiple foci of telomeric

DNA and interestingly CtIP colocalised with some of the telomeric DNA foci in all ALT cell lines (Figure 5.1). The percentage of cells with CtIP/DNA colocalization foci ranged from 11.7% to 38%, among the ALT+ cell lines (Table 5.1 and Figure 5.2), whereas CtIP/DNA colocalization foci in telomerase+ cells were found in just 0.08% of the HT1080 cells and none of the HeLa cells. These findings show that CtIP is associated more frequently with telomeric DNA in ALT cell lines than in telomerase+, this association could be at telomeres or with telomeric extrachromosomal DNA or both.

Table 5.1 Cell lines were screened for colocalization of CtIP and telomeric DNA. All the cells were positive for telomeric DNA signal. The last column shows the total number of cells screened. Numbers in parenthesis represents the number of positive cells for colocalization foci (CtIP/DNA) or no colocalization (CtIP-only foci).

Cell line	Percentage of positive cells for foci of CtIP/DNA colocalization (number of cells)	Total cells scored
HT1080	0.08% (1)	1155
HeLa	0	1274
WI38V13/2RA	38% (42)	110
SUSMI	23% (25)	107
W-V	25% (29)	115
SAOS	11.7% (12)	114

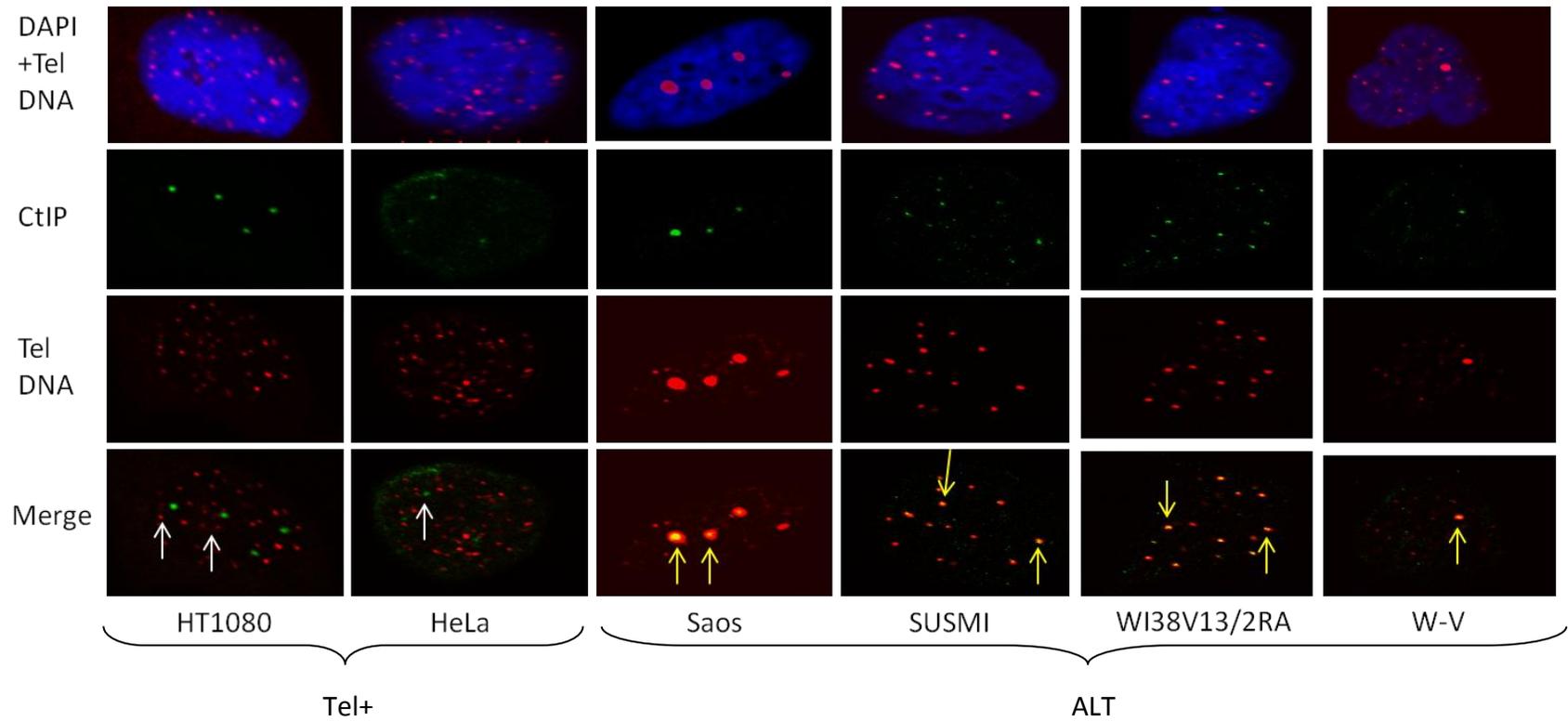


Figure 5.2 Immunofluorescence in the telomerase+ cell lines HT1080 and HeLa and the ALT+ cell lines SAOS, SUSMI, WI38V13/2RA and W-V using a polyclonal antibody against CtIP (green 2nd row) and a PNA Telomeric probe (red 3rd row). The bottom row shows the merged images and the white arrows point to some CtIP foci without colocalization to telomeric DNA in the telomerase+ cell lines. The yellow arrows point to some of the foci of colocalization in the ALT cell lines. Images were captured with a Olympus confocal FV1000 microscope.

5.4 Investigation of CtIP in ALT positive cell lines

5.4.1 Colony forming assay in ALT cells transfected with a vector expressing a shRNA against CtIP

To investigate whether downregulation of CtIP has an effect on telomere function in ALT positive cell lines, two shRNAs against *CtIP* were designed: shCtIPA (sequence published (Huertas P *et al.* 2009) and shCtIPB (designed with siRNA target finder tool from Ambion). WI38V13/2RA and SUSM1 cells were transfected with the pSUPERIOR vector containing shCtIPA, shCtIPB or an empty vector as control. The WI38V13/2RA cells were transfected using the FuGENE reagent (Promega) and the SUSM1 cells were transfected by electroporation (Gene Pulser Xcell™ electroporation system Bio-Rad). For the SUSM1 cell line three million cells were used per electroporation per plasmid. Electroporated cells were divided among three 10cm Petri dishes with 15ml of medium. For the WI38V13/2RA one million cells were plated in 10cm petri dishes in 15ml of medium 24 hrs before transfection. Each cell line was transfected with the shCtIPA, shCtIPB or shControl in duplicate. After transfection cells were cultured in media with G418 (500µg/ml) to select for cells that contained the pSUPERIOR vector.

To evaluate whether downregulation of *CtIP* affected colony forming efficiency in ALT cell lines the transfected cells were growth for 14 days and then plates were fixed and stained with 1% crystal violet and colonies were counted. In the WI38V13/2RA cell line the number of colonies counted were 33, 36 and 5 for the shControl, shCtIPA and shCtIPB respectively, and for the SUSM1 there were 8, 3 and 1 for the shControl, shCtIPA and shCtIPB respectively (Figure 5.3). The shCtIPA seemed to have no effect on colony formation in the WI38V13/2RA cells and a mild effect on the SUSM1 cells. On the other hand shCtIPB reduced the colony formation efficiency in both cell lines. To investigate further whether this effect has caused by *CtIP* downregulation in ALT cells, clones stable expressing the shCtIP needed to be generated.

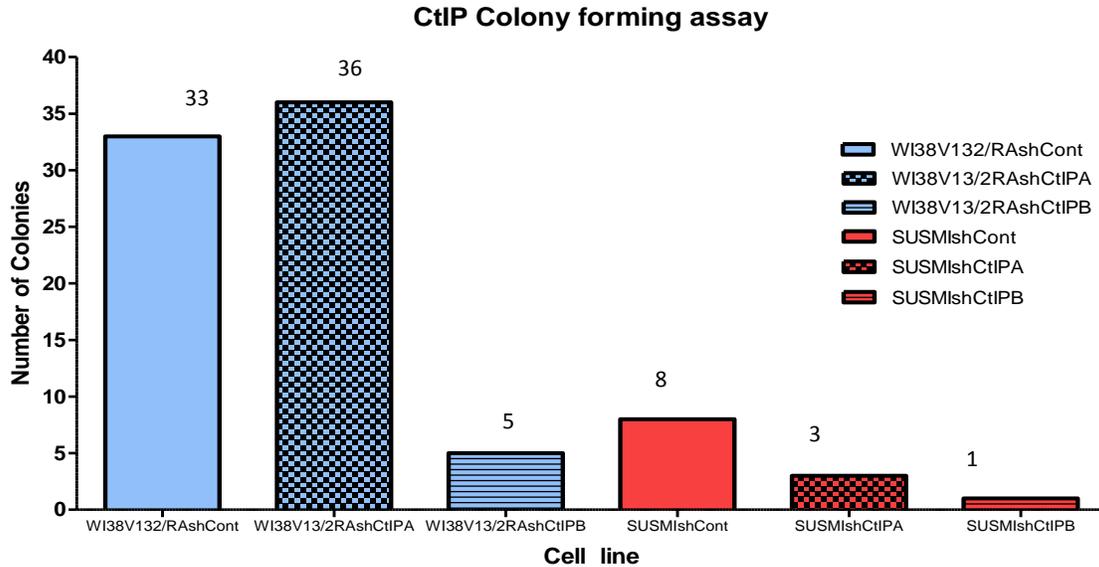


Figure 5.3 Colony forming assay following transfection with shCtIPA or shCtIPB plasmid. The WI38V13/2RA (one million cells per transfection) and SUSM1 (three million cells per transfection) ALT cell lines were transfected with the pSUPERIOR vector containing a shCtIPA or shCtIPB or an empty vector as control.

5.4.2 Expansion of clones transfected with a vector expressing shCtIPA, shCtIPB or shControl

The ALT cell lines WI38V13/2RA and SUSMI were transfected as described above. After transfection cells were cultured in media with G418 (500 μ g/ml) to select for cells that contained the pCtIP construct. Once the colonies were visible they were recovered and expanded. The number of population doublings (PDs) was recorded at each passage and growth curves were generated (Figure 5.4). For the WI38V13/2RA cell line four control clones and one shCtIPA clone were recovered, but no clones were recovered following transfection with the shCtIPB. The control clones showed similar growth rates reaching PD20 approximately at day 34, while the shCtIPA1 clone grew more slowly and reached the same point at day 46 (Figure 5.4A). From SUSM1 five control clones, two shCtIPA and two shCtIPB clones were recovered. The growth curves show that the control clones grew at similar rates reaching PD20 at around day 35. One of the shCtIPB clones (B1) showed similar growth rate to control clones suggesting there was no downregulation of *CtIP*. The

other three (A1, A3 and B2) showed slower growth rates reaching PD20 after day 60 (Figure 5.4B). RNA was extracted at ~PD20 and after cDNA synthesis qPCR was conducted to determine whether *CtIP* expression was reduced and to check if there was a relationship between the growth rates and *CtIP* downregulation.

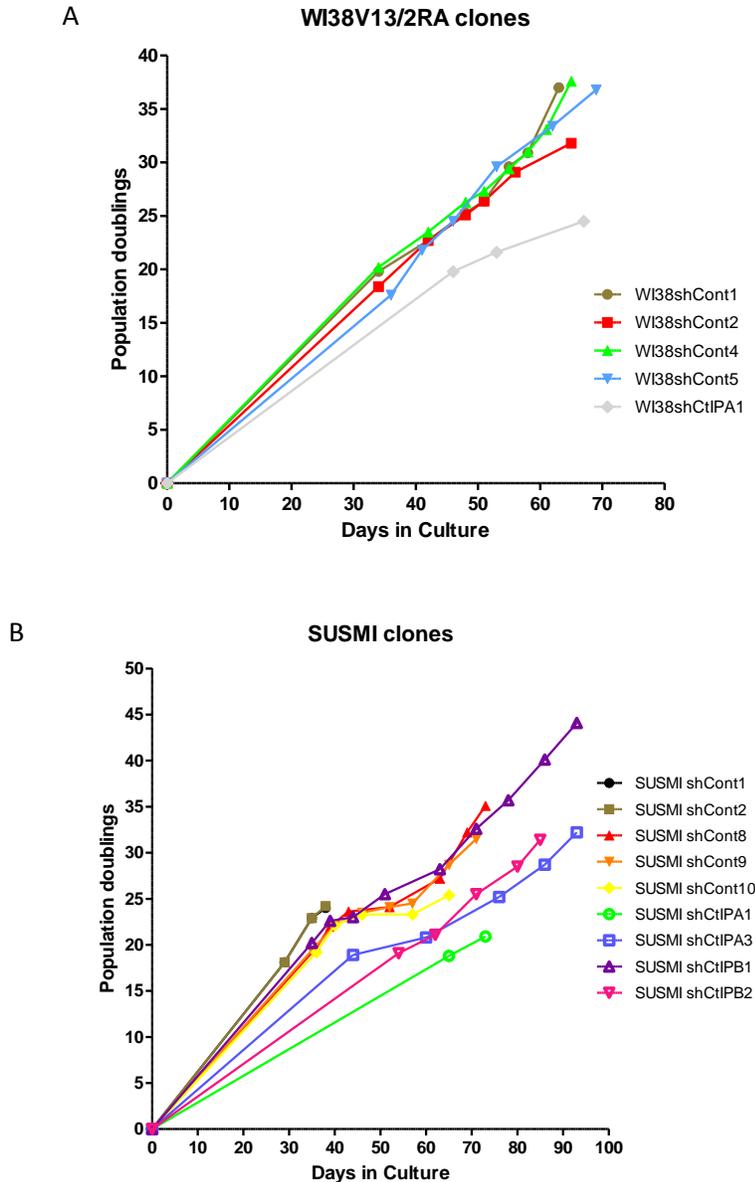
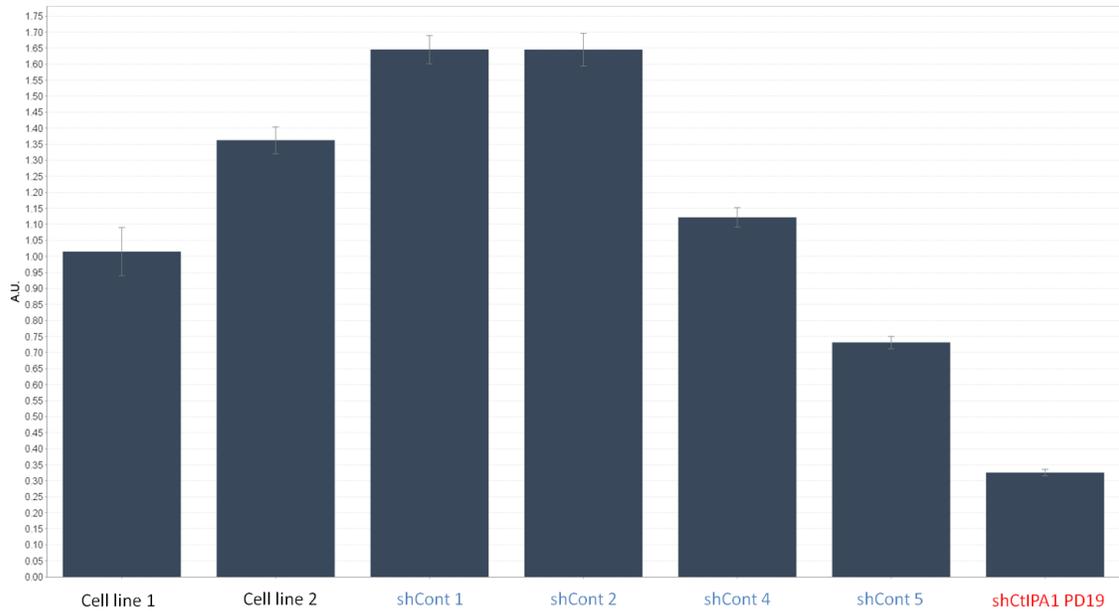


Figure 5.4 Growth curves of the expanded clones from the ALT cell lines WI38V13/2RA (A) and SUSMI (B) after transfection with a shControl, shCtIPA or shCtIPB plasmids.

For the qPCR analysis two independent cDNA syntheses from the progenitor cell lines were included as controls. The *CtIP* expression analysis in control clones derived from the WI38V1372RA cell line showed clonal variation, as observed in the shEXO1 experiments. In spite of this the clone shCtIPA1 showed ~50% reduced expression of *CtIP* expression at PD19 when compared with the control clone with the lowest expression (shControl5) (Figure 5.5A). The experiment was repeated including cDNA from clone shCtIPA1 at PD19 and PD21. At both PDs the level of downregulation observed was basically identical confirming at least 50% downregulation of *CtIP* expression (Figure 5.5B), suggesting that the delayed growth observed in this clone could be related to the reduced amount of *CtIP* expressed.

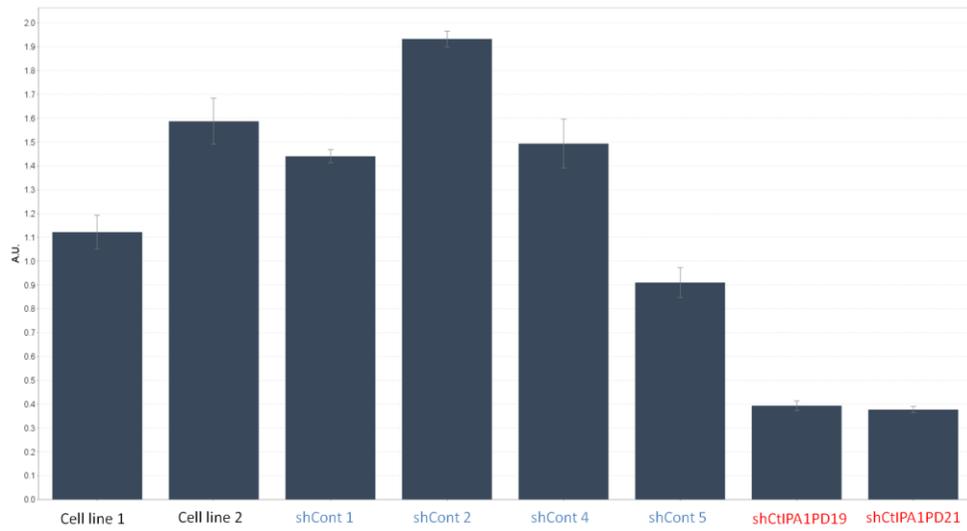
For the SUSMI cell line the variation observed in CtIP expression between control clones made it difficult to determine whether there was a real effect on CtIP expression in the presence of shCtIPA or B. For example shCtIPA1 and shCtIPA3 seem to have low expression of *CtIP* but shControl9 showed even lower levels of expression (Figure 5.5C), the experiment was repeated on the same cDNA but similar results were obtained (see appendix 5). In summary none of the SUSMI shCtIP clones was considered to show significant reduction of *CtIP* expression.

A



WI38V13/2RA

B



WIA38V13/2RA

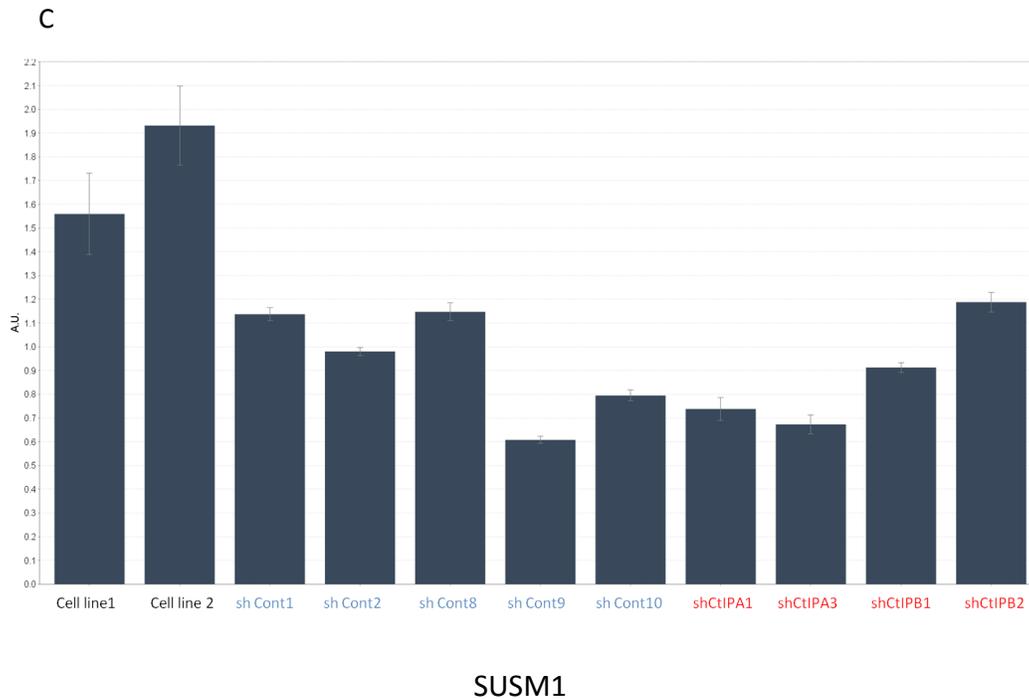


Figure 5.5 Expression analysis of CtIP by qPCR in the clones recovered from the ALT+ cell lines after transfection with a vector expressing a shRNA against *CtIP* (shCtIPA or shCtIPB), or an empty vector (shCont). Two independent cDNA syntheses from the respective progenitor cell line were included in the analysis (Cell line 1 and Cell line 2). The WI38V13/2RA clone shCtIPA1 showed ~50% downregulation of *CtIP* expression at PD19 (A) analysis was repeated on the same clone at PD21 showing the same level of downregulation (B). None of the SUSM1 shCtIP clones showed a significant reduction of *CtIP* expression. For technical replica see appendix 5 (C). All the reaction were done in triplicate. A.U. Arbitrary Units.

5.5 Telomere analysis in a WI38V13/2RA clone shCtIPA (with downregulation of *CtIP*) using the STELA method

Telomere length in ALT+ cells is known to be very heterogeneous with molecules up to 50kb, and because of this feature the telomere restriction fragment (TRF) assay is regarded as a reliable technique to detect the longest telomere molecules, and to look for changes in telomere length. However, STELA profiles can reveal changes in the proportion of short telomeres that can be missed by TRF analysis. In this case both techniques were used to investigate the effect of *CtIP* downregulation on telomere length. DNA from the

WI38V13/2RA clone shCtIPA1 was prepared for STELA as described in the methods section. Analysis of the XpYp telomere length showed that the median telomere length from the shCtIPA1 clone (698.4 bp) was significantly shorter (Mann Whitney $p < .0001$) than in four control clones (Figure 5.6). In addition there was no evidence of length heterogeneity as seen in the control clones. To determine whether this was an effect of *CtIP* downregulation the 12q and 17p telomeres were investigated. STELA analysis showed that the 12q and 17p telomeres in the shCtIPA1 clone were not the shortest among the clones analysed (Figure 5.7). Furthermore all clones revealed much greater telomere length heterogeneity, as expected for ALT+ cell lines (Southern blots are shown in Appendix 6 section A6.1).

To investigate if there was an effect on overall telomere length TRF analysis also was conducted on all the clones and DNA from the progenitor cell line was included. The analysis did not reveal any significant differences in the median from the shCtIPA1 clone compared to controls and the progenitor cell line. In addition, no effect was seen on total amount of telomeric DNA since the intensity of the signal is very similar among the samples analysed. Following hybridisation to the telomere probe the Southern blot was hybridised to the MS1 probe as a loading control. MS1 detects a hypervariable single copy minisatellite within the human genome (Figure 5.8). These findings show that ~50% downregulation of expression of *CtIP* did not have a significant impact on telomere length or abundance of telomeric DNA in the WI38V13/2RA shCtIPA1 clone. The short XpYp telomere observed in the shCtIPA1 clone may have arisen by chance at this chromosome end in the cell that seeded the clone.

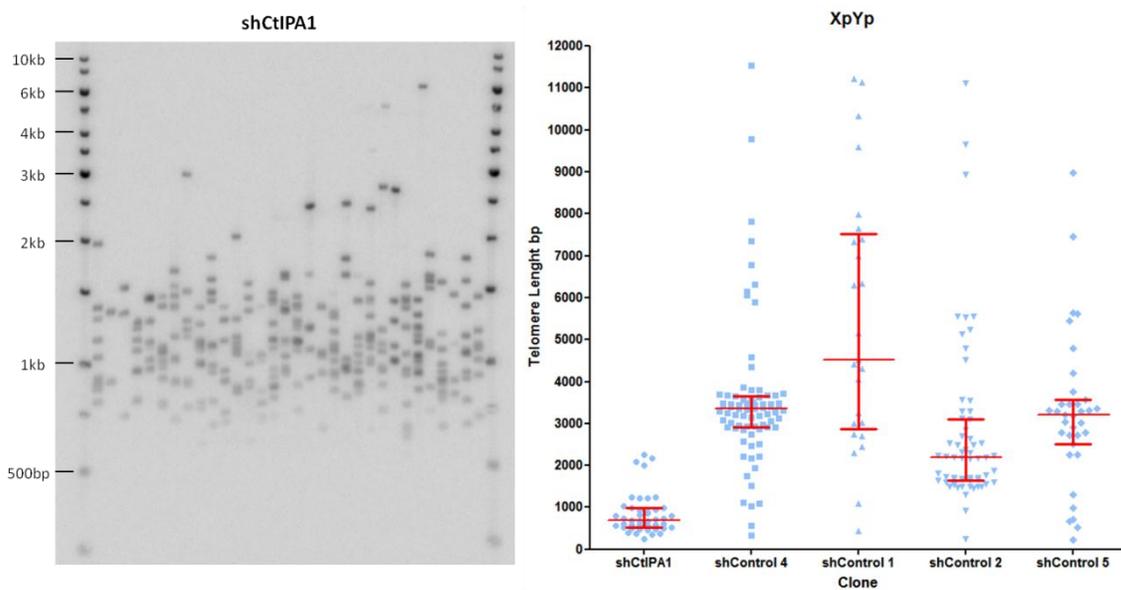


Figure 5.6 Length analysis by STELA of the XpYp telomere of clone shCtIPA1 and control clones, at PD19, after adjustment for flanking DNA in each STELA product (415 bp for XpYp). On the left a representative Southern blot of XpYp STELA from the shCtIPA1 clone. Additional Southern blots were done to complete the analysis. On the right comparison of the XpYp telomere analysis between the shCtIPA1 clone and four control clones showing the median and interquartile range. The difference between the shCtIPA1 clones and controls is statistically significant (Mann Witney $p < .0001$).

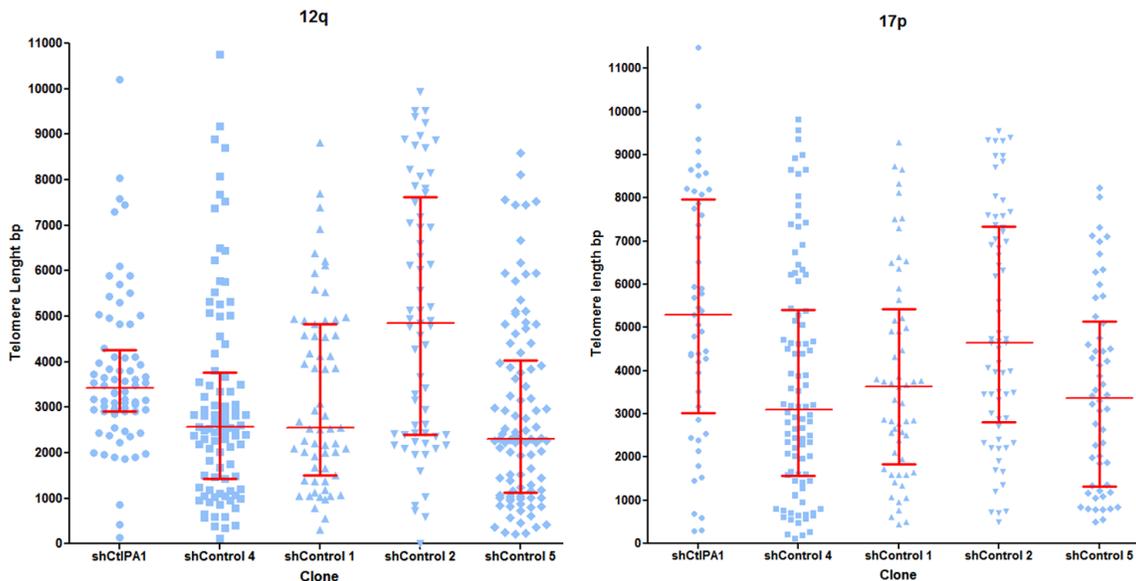


Figure 5.7 Analysis of STELA products from telomeres 12q and 17p obtained from the shCtIPA1 and four control clones after adjustment for flanking DNA in each STELA product (532 bp for 12q and 3078 bp for 17p). Bars indicate median and

interquartile range. The heterogeneous length distribution is expected in ALT+ cell lines.

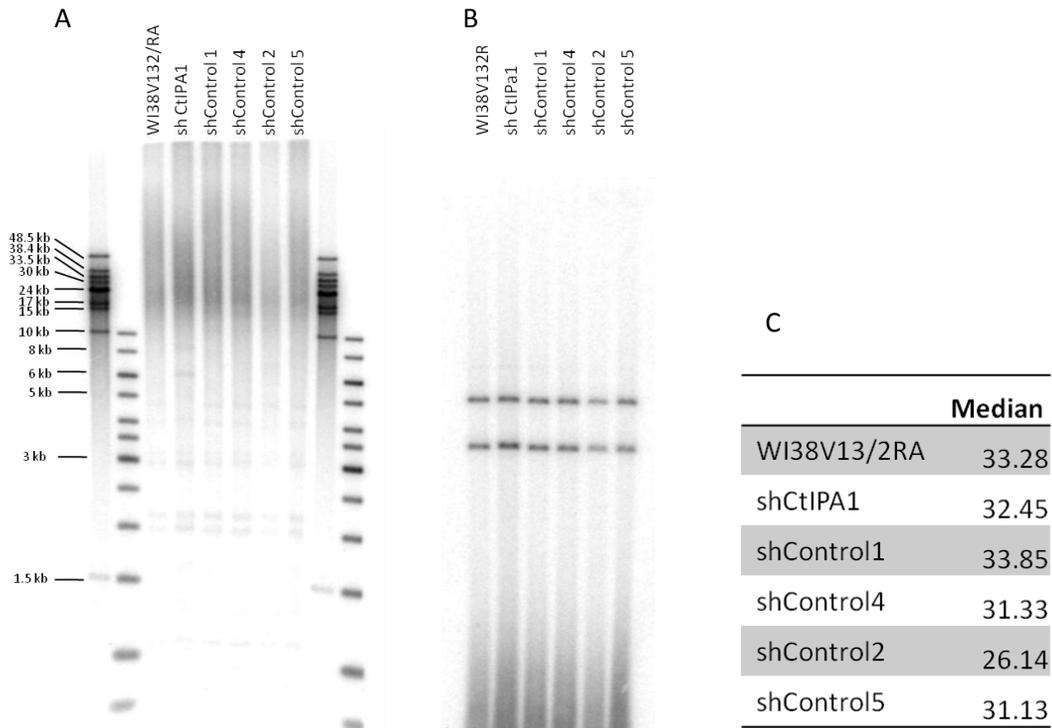


Figure 5.8 TRF of WI38V13/2RA cell line and clones. (A) Southern blot hybridization of 4 μ g of genomic DNA digested with *HinfI* and *RsaI* and hybridised to a 32 P labelled telomere probe. (B) Hybridisation of the same Southern blot to the MS1 probe as a loading control. (C) Median telomere length determined from A using the telometric software.

5.6 Effect of the downregulation of CtIP in the shCtIPA1 clone on markers of ALT activity

5.6.1 Effect of downregulation of CtIP on the MS32 minisatellite

The instability of the MS32 minisatellite is a known feature of ALT cell lines (Jeyapalan JN *et al.* 2005), and downregulation of the BLM helicase lead to an increase in MS32 instability in the ALT cell line W-V (Mendez-Bermudez *et al.*, unpublished data). To investigate whether 50% reduction of the *CtIP* expression affected MS32 instability, mutation rates in the shCtIPA1 and control clones were measured using SP-PCR (see methods). The mutation rate calculated for the shCtIPA1 clone was 0.00132/PD and the average mutation rate for the control clones was 0.00072/PD. The difference is not

statistically significant ($p=0.57$ Mann Whitney) showing that the level of *CtIP* downregulation had no effect on MS32 instability. Figure 5.9 shows the southern blot used to score mutations in the shCtIPA1 clone and the southern blot used for Poisson analysis.

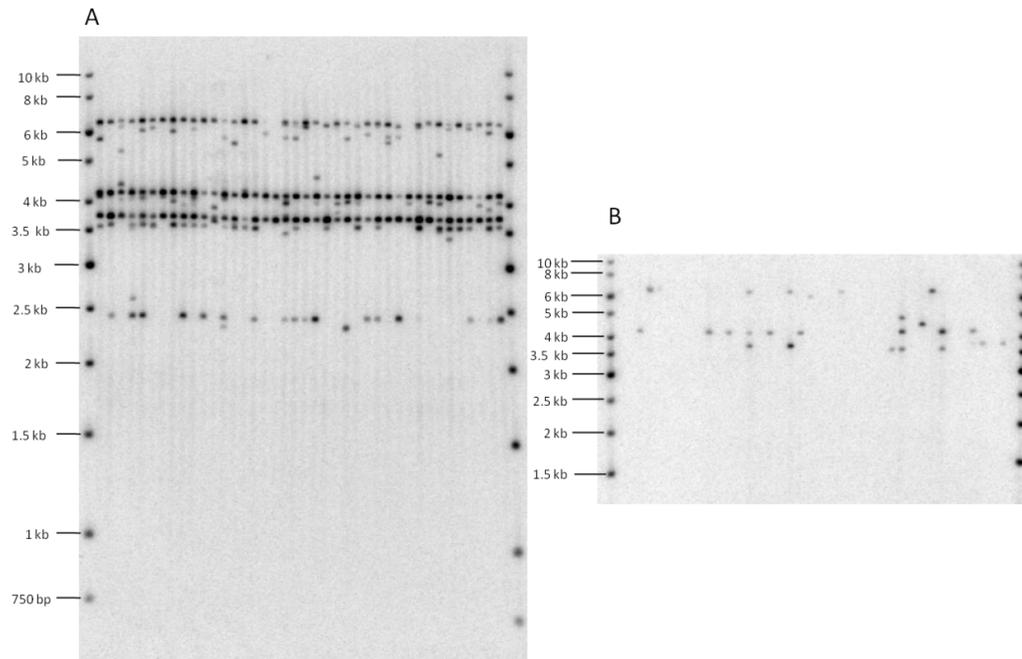


Figure 5.9 Southern blots of the gels used to calculate MS32 mutation rate on the WIV13/2RA shCtIPA1 clone. (A) For small pool PCR (SP-PCR) 40 reactions were prepared to amplify the MS32 minisatellite, each reaction contained 100 pg of genomic DNA digested with *MboI*. The W138V13/2RA cell line has three copies of chromosome 1 hence it has three progenitor alleles (~3.7 kb, ~4.2 kb and ~6.2 kb). Bands different in length to the progenitor alleles are mutant molecules. When the number of mutants was counted, those mutants seen more than three times were considered as derived from the same mutation event at early stages of clonal expansion and therefore were scored just once. To calculate the total number of molecules screened by SP-PCR and obtain the mutation frequency and rate Poisson analysis was conducted (B). For Poisson analysis 40 PCR reactions each containing an estimated 5 pg of genomic DNA digested with *MboI* were prepared. The number of positive wells was used to calculate the number of amplifiable molecules per track of the SP-PCR and then the total number of molecules screened.

5.6.2 Effect of downregulation of CtIP on the C-circle formation

In a recent study, C-circles, which are partially single stranded circles of extrachromosomal telomeric DNA with a continuous C-rich strand, have been proposed as the most reliable marker of ALT activity. This is because C-circles have been detected even in the absence of other ALT markers such as the APBs (Henson JD *et al.* 2009). The shCtIPA1 clone was analysed to look for a change in the level of C-circle DNA as a consequence of *CtIP* downregulation. The C-circle assay is a DNA amplification method to detect circular, partially single stranded, extrachromosomal telomeric DNA. The method was described by Henson JD *et al.*, (Henson JD *et al.* 2009) and it takes advantage of the properties of the DNA polymerase from the phage phi 29 (Φ 29). Fragments of the discontinuous G-rich strand present in the C-circles act as primers for the DNA synthesis. The synthesis assay does not contain dCTP to ensure the amplification exclusively of C-circles. The level of C-circle DNA was not affected by reduction of *CtIP* expression (Figure 5.10). The C-circle assay was repeated on the same DNA preparations and similar results were obtained (Appendix 6 section A6.2).

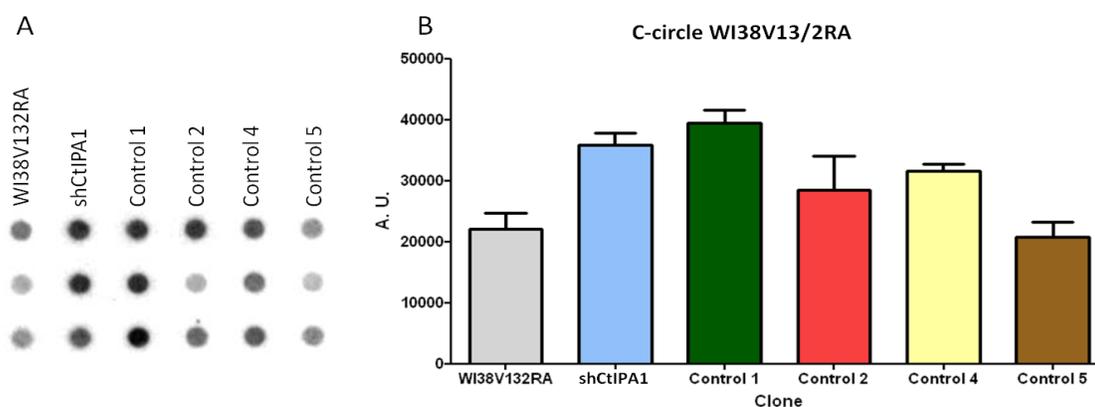


Figure 5.10 Quantification of telomeric DNA generated in the C-circle assay from the WI38V13/2RA cell line, shCtIPA1 and control clones. (A) Dot blot of the C-circle amplification reactions after hybridisation to a ^{32}P labelled telomere probe. Reactions were done in triplicate. Each reaction contained 10 ng of *EcoRI* digested genomic DNA. (B) Signal intensities were analysed using the Dot blot analyser from ImageJ software to generate plots. A.U. Arbitrary Units.

5.6.3 Effect of downregulation of *CtIP* on telomere mutation

A mutation frequency of ~5.5% has been observed in ALT cells, most of them being complex mutations (Varley H *et al.* 2002). To investigate the effect of the *CtIP* downregulation on telomere mutation frequency in the shCtIPA1 clone, single molecule STELA analysis was combined with TVR-PCR to identify mutant molecules. First the XpYp telomere maps in the progenitor cell line were determined (Figure 5.11). Then 122 single STELA products from the shCtIPA1 clone were generated and analysed by TVR mapping. All the mapped molecules corresponded to map 1 (Figure 5.11) and no mutations were observed (Appendix 6 section A6.3). The finding that all the mapped molecules corresponded to map 1 may suggest the loss of one sex chromosome in the shCtIP clone. The absence of mutations was unexpected as previous analysis of the WI38V13/2RA cell line showed a mutation frequency of 5.5% (Varley H *et al.* 2002). One possible explanation is that the XpYp telomeres analysed show relatively few variant repeats (Figure 5.11) and so mutations may go undetected. The shortness of the XpYp telomere (Figure 5.6) suggests this telomere may not been engaged in telomere-telomere recombination. Attempts to investigate the telomere mutation frequency at other telomeres in the shCtIPA1 clone were made. The 12q telomere was chosen because it has been described that the WI38V13/RA cell line possesses a polymorphism in the 12q subtelomeric region. This polymorphism consists of a ~1kb deletion which generates a unique sequence that can be used to amplify the associated telomere specifically; this allele is called 12q Δ . Attempts to amplify the 12q Δ allele with a specific primer for the 12q Δ allele were made (Baird DM *et al.* 2000). No STELA products were obtained from the clone shCtIPA1 and the fact that STELA products were obtained using DNA from the progenitor cell line as positive control suggests that the shCtIPA1 clone has lost the 12q Δ allele (Figure 5.12). An attempt was made to map the distribution of variant repeats at the 17p telomere, but as in XpYp the pattern of repeats did not show much variation which makes the detection of mutants difficult. The lack of an informative TVR map prevented further analysis of telomere mutations in the shCtIPA1 clone.

Map 1 GGGGTTGGT(41)CCCT...
 Map 2 GGGGGGTTGGTTTGGT...
 T= TTAGGG G=TGAGGG C=TCAGGG

Figure 5.11 XpYp telomere maps for the WI38V13/2RA cell line. The WI38V13/2RA cell line is homozygous for the SNPs located in the subtelomeric region at 415 and 30 bp upstream of the first telomeric repeat (-415C/C and -30T/T). The primer used for STELA is specific for the -415C allele and the primer used for the TVR-PCR is a nested primer specific for the -30T allele. Dots mean continuation of the TTAGGG repeats.

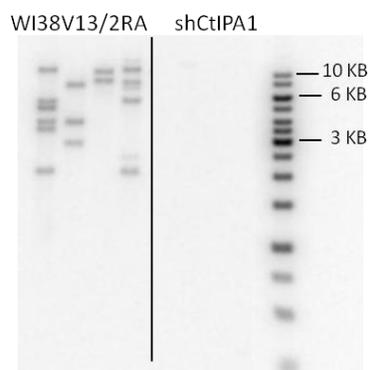


Figure 5.12 Southern blot of the 12q Δ STELA in the WI38V13/2RA cell line and the shCtIPA1 clone. The 12q Δ allele was amplified using a specific primer 12qNull3 described previously (Baird DM *et al.* 2000). The 12q Δ allele was effectively amplified in the WI38V13/2RA cell line but not in the shCtIPA1 clone. Each PCR contained 250 pg of *EcoRI* digested genomic DNA. The membrane was hybridised with a ^{32}P labelled telomere probe.

5.6.4 Detection of the telomeric 3' single strand overhang in WI38V13/2RA cell line

CtIP is thought to take part in the formation of the 3' overhang at telomeres (Bonetti D *et al.* 2009, Iglesias N *et al.* 2009), so the next aim was to investigate changes in the total amount of the overhang after downregulation of *CtIP*. In 2007 Zhao Y *et al.*, described a method to measure the length and total amount of 3' single stranded telomeric overhang taking advantage of a specific double strand nuclease (DSN) that specifically degrades specifically dsDNA to fragments $\sim 10\text{bp}$, leaving intact ssDNA. Following digestion with DSN

the single stranded DNA was size separated in denaturing polyacrylamide gels, transferred to a nylon membrane and hybridized with a radiolabelled (CCCTAA)₃ oligonucleotide probe, (Zhao Y *et al.* 2007). The technique seemed appropriate for detecting changes to the length or amount of 3' single stranded telomeric overhang after *CtIP* downregulation. For optimization genomic DNA from ALT and telomerase+ cell lines was used. Genomic DNA (5µg) was digested with the DSN, and as a negative control another 5µg of genomic DNA were incubated for 1hr with *E. coli ExoI* to degrade ssDNA, before digestion with DSN. One modification to the protocol was the use of a double stranded telomere probe in which only the C strand is radiolabelled with ³²P. The first attempts showed very weak signal and also there was signal in the negative control. The possible explanation for having signal in the negative control was that the *ExoI* buffer was inhibiting the DSN activity. To avoid this, the DNA was purified with phenol/chloroform/IAA and ethanol precipitation after incubation with *ExoI* and prior to DSN digestion. New attempts showed very weak signal in the positive samples (DSN digested only), this could be due to inefficient transfer from the gel to the membrane. To avoid loss of ssDNA during transfer, samples were attached to a nylon membrane by dot blot. By doing this it was not possible to measure the length of the ssDNA, but it was still possible to assess intensity of the signal. In one of the dot blots some signal was detected in the samples treated with DSN but not in those previously incubated with *ExoI*. The probe was removed and the membrane was hybridized with a chromosome 7 alpha-satellite probe (D7Z1), no positive signal was observed which suggests that the positive signal with the telomere probe was not just a non-specific signal (Figure 5.13). In spite of having a positive result there are two complications with the technique. First at least 10µg of genomic DNA are needed and second if the method is going to be used to compare between clones there should be a control to guarantee that initial amount of genomic DNA to be treated is identical between all the samples.

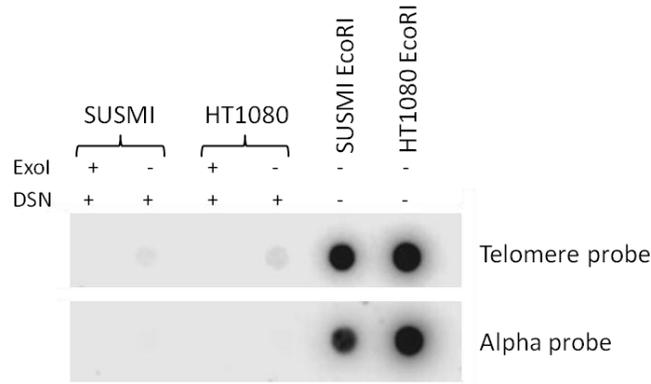


Figure 5.13 Detection of single stranded (TTAGGG)_n DNA by dot blot. On top hybridization with the telomere probe, positive signal was observed in samples treated only with DSN not in the samples previously incubated with Exo1. On the bottom after probe removal the membrane was hybridised with a chromosome 7 alpha-satellite probe (D7Z1), no signal was detected. As positive controls for hybridization conditions 100ng of *EcoRI* digested DNA from each cell line were used.

5.7 Effect of transient downregulation of *CtIP* on telomere length and C-circle formation on ALT cell lines

No changes were observed in the ALT markers studied in the shCtIP1 clone after 19PD, and so to answer whether downregulation of the *CtIP* has an immediate effect on ALT the cell lines were transfected with a siRNA against *CtIP* (siCtIP) or a scrambled siRNA as a control (siControl). To check for effective downregulation the telomerase+ HT1080 and the ALT cell lines WI38V13/2RA and SUSM1 were assessed by qPCR for expression of *CtIP*. GAPDH, β -Actin and UBC (Primer Design) were used as reference genes and *CtIP* expression measured 48 hrs after transfection. The qPCR was done twice on the same cDNA and this showed the same results (Figure 5.14 and appendix 6 section A6.4). The level of downregulation achieved in HT1080, SUSM1 and WI38V13/2RA was 85%, 40% and 80% respectively. Telomere length was analysed by TRF on all samples and the median was calculated using the telometric software. Although the telomere length of the SUSM1 and WI38V13/2RA cell lines treated with siCtIP suggested there has been some shortening compared to controls, hybridisation to the MS1 probe showed that the amount of DNA

loaded into the gel was uneven between pairs of tracks (Figure 5.15). Therefore the data suggest telomere length was not significantly affected by 48 hrs of *CtIP* downregulation.

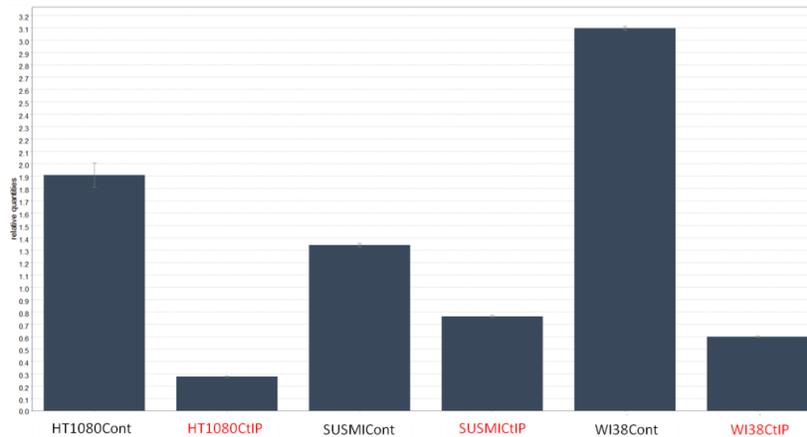


Figure 5.14 qPCR analysis of *CtIP* expression in HT1080, SUSM1 and WI38V13/2RA 48hrs after transfection with siCtIP or siControl. The downregulation achieved was 85%, 40% and 80% respectively. Data was normalised using GAPDH, β -Actin and UBC as reference genes. The reactions were done in triplicate and data analysed using the qbase^{plus} software (Biogazelle)

Quantification of C-circle DNA was done on the ALT cell lines SUSM1 and WI38V13/2RA. The analysis showed that circles were significantly reduced in both cell lines, In WI38V13/2RA a twofold reduction and in SUSM1 approximately a 10 fold reduction was observed. The experiment was repeated using the same DNA preparations and the difference in C-circle abundance in WI38V13/2RA was the same, but in SUSM1 the difference was not as striking, showing a twofold reduction, that was significant (Figure 5.16). This effect was not seen in the clone shCtIPA1 at PD19 suggesting that the loss of C-circles could be a short term effect. The telomerase+ cell line HT1080 was not include in the analysis since Henson JD *et al*, reported, (after studying 15 telomerase+ cell lines), that the C-circle DNA is practically undetectable in telomerase+ cells (Henson JD *et al*. 2009).

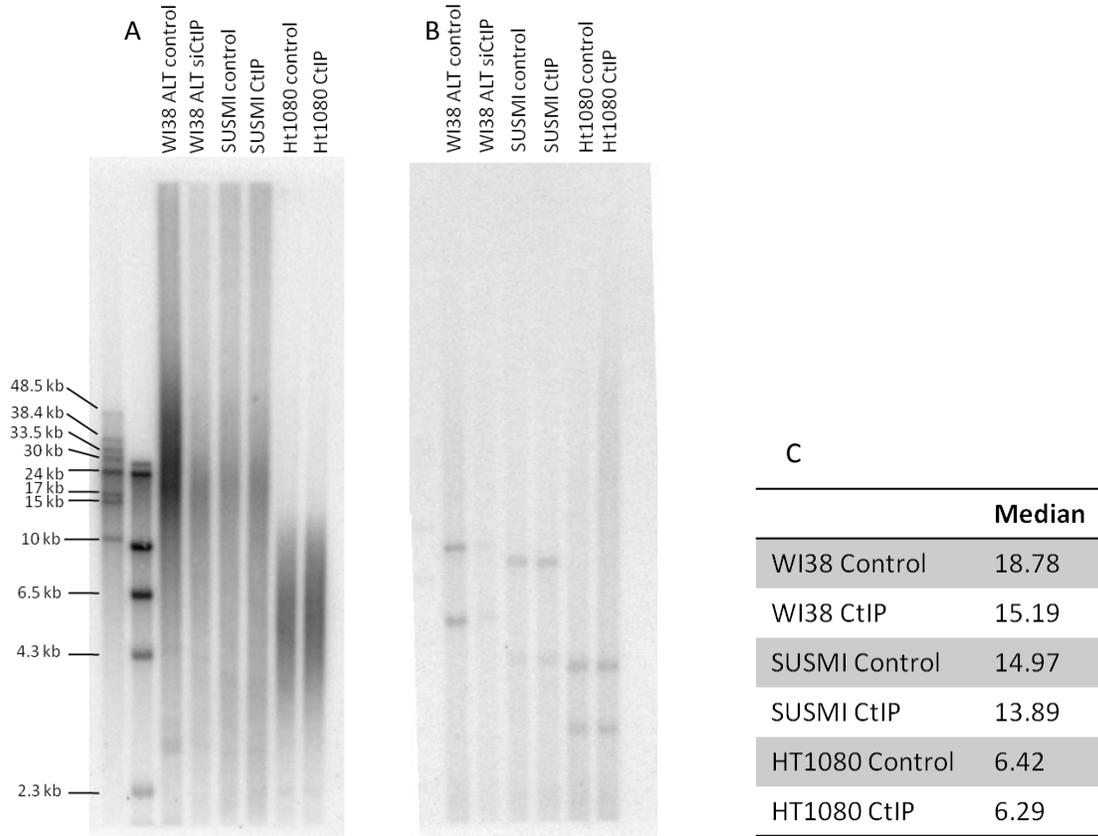


Figure 5.15 TRF of ALT+ WI38V13/2RA, SUSM1 and the telomerase+ HT1080 cell lines. (A) Southern blot hybridization of 4 µg of genomic DNA digested with *Hinfl* and *RsaI* and hybridised to a ³²P labelled telomere probe. (B) Hybridisation of the same Southern blot to the MS1 probe as a loading control. (C) Median telomere length determined from A using the telometric software.

In order to investigate further the effect of *CtIP* expression on the presence of C-circles the transfection was repeated in the same cell lines and in W-V and SAOS ALT cell lines, and the telomerase+ cell line HeLa. Quantification of *CtIP* expression using qPCR was done 48hrs after transfection and the levels of downregulation achieved were HT1080 85%, HeLa 75%, SAOS 80%, SUSMI 50%, WI38V13/2RA 75% and W-V 70%. The qPCR analysis was repeated on the same cDNA showing the same results (Figure 5.17 and appendix 6 section A6.4). The reduction in *CtIP* expression was very similar between both transfections, confirming that the siCtIP gives constant results. Interestingly the SUSMI cell line, which provided the only clone with reduced *CtIP* expression upon transfection

with the shCtIP vector, showed the least *CtIP* downregulation, this may be caused by a less efficient transfection.

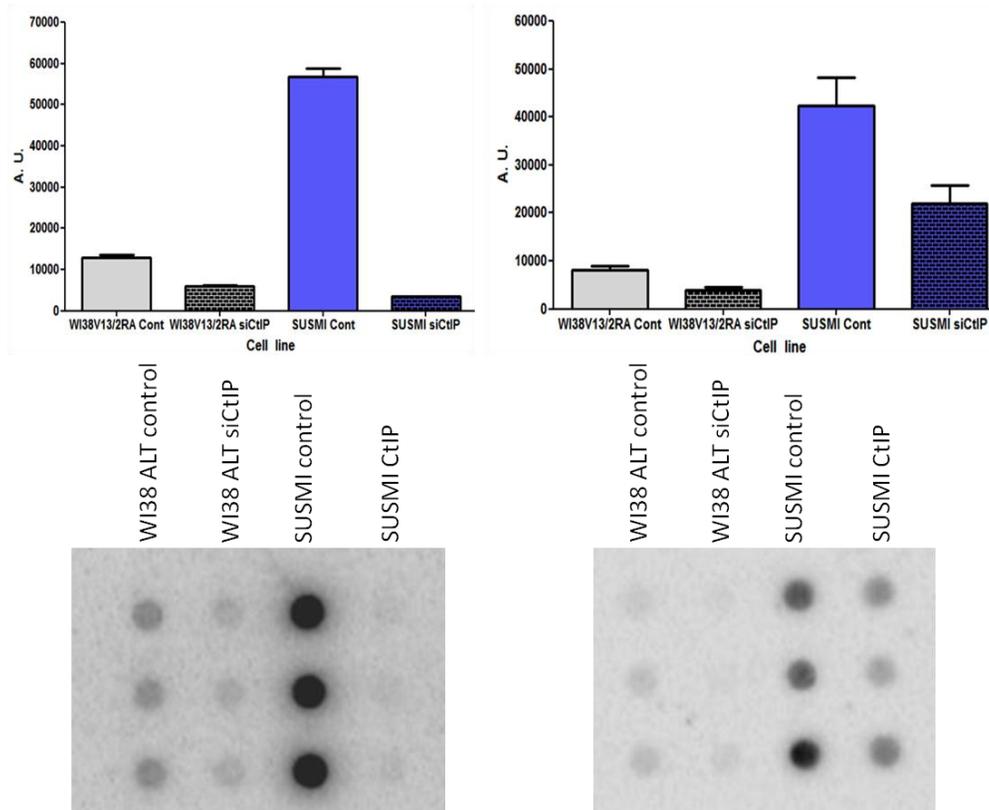


Figure 5.16 Quantification of the C-circle detection assay on WI38V13/2RA and SUSM1 cell lines 48hrs after transfection with siCtIP or siControl. The first assay (left) showed a twofold reduction in WI38V13/2RA ($p=0.005$ Student's *t* test) and a 10 fold reduction in SUSM1 transfected with siCtIP ($p < 0.0001$ Student's *t* test). The replica of the experiment (right) showed the same pattern in WI38V13/2RA ($p=0.01$ Student's *t* test) but in SUSM1 the reduction was smaller, though is still significant ($p=0.04$ Student's *t* test). Dot blots after hybridisation with a ³²P radiolabelled telomere probe are shown below each plot. Signal intensities were measured with dot blot analyser from ImageJ software and data was analysed using GraphPad Prism 5 software. A.U. Arbitrary Units.

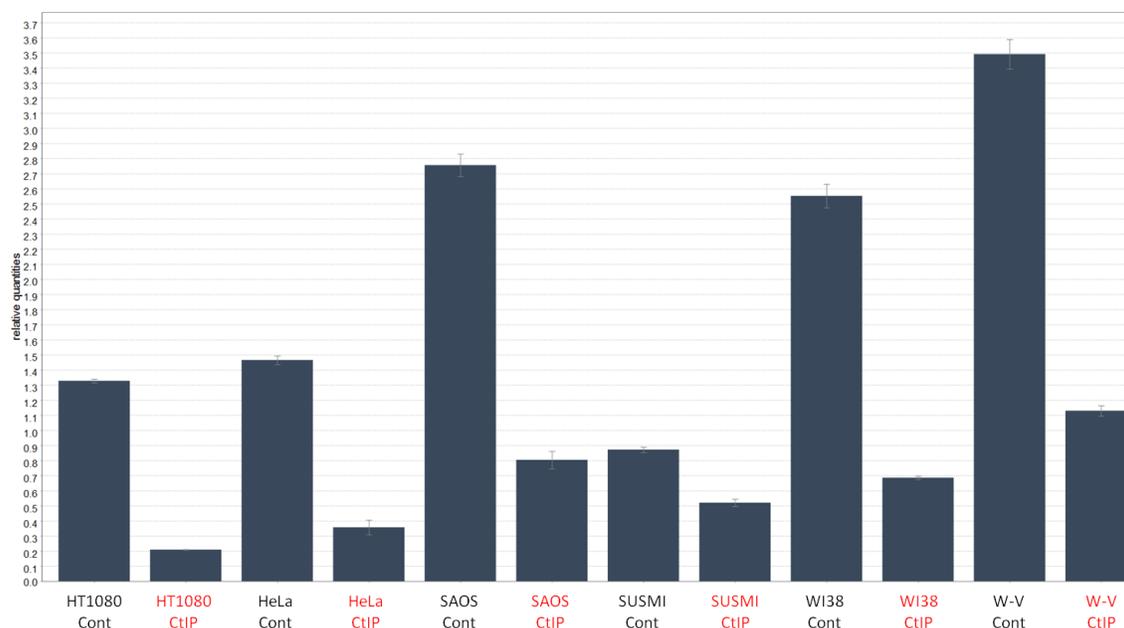


Figure 5.17 qPCR analysis of *CtIP* expression in the HT1080, HeLa, SAOS, SUSM1, WI38V13/2RA and W-V 48hrs after transfection with siCtIP or siControl. Data was normalised using GAPDH, β -Actin and UBC as reference genes. The downregulation achieved was 85%, 75%, 80%, 50%, 75% and 70% respectively. Reactions were done in triplicate. Data was analysed using qbase^{plus} software (Biogazelle). Replica of the qPCR analysis was done on the same cDNA showing the same results (see Appendix 3)

Unexpectedly the C-circle assay did not give reproducible results as there was no reduction in C-circle abundance following the second transient downregulation of *CtIP* in the WI38V13/2RA or SUSM1 cell lines. The results showed no change in C-circle abundance in the WI38V13/2RA cell line. However, the rest of the ALT+ cell lines showed an increase in the levels of circles after siCtIP transfection, especially in the W-V cell line (Figure 5.18). The C-circle assay was repeated on the same DNA preparations and similar results were found (Appendix 6 section A6.5). The chance of degradation of circles during DNA extraction seems unlikely because this would imply loss of C-circle DNA only in the control cells of the three ALT+ cell lines studied.

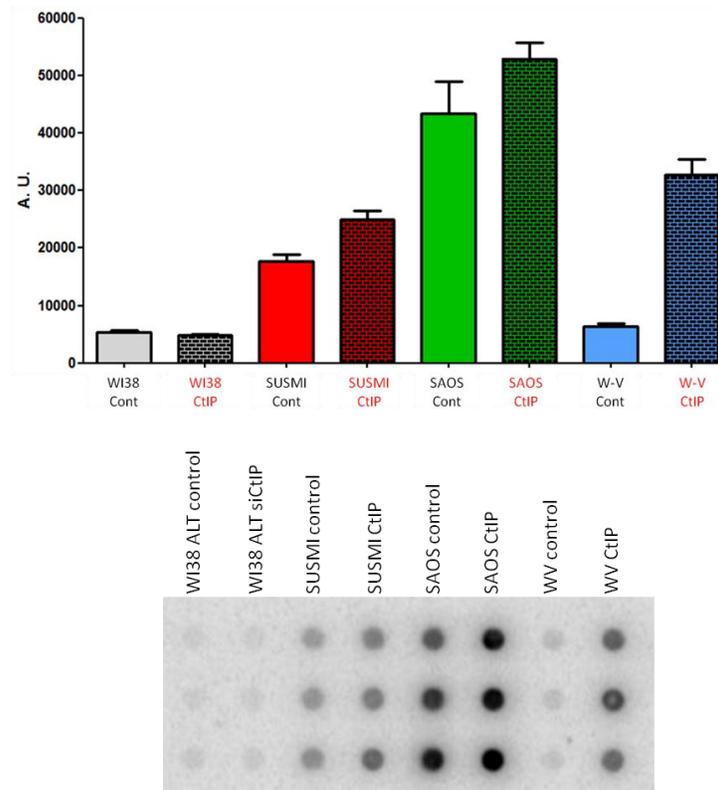


Figure 5.18 Quantification of C-circle DNA in WI38V13/2RA, SUSM1, SAOS and W-V ALT+ cell lines 48hrs after the second transfection with siCtIP or siControl. The C-circle abundance did not change in the WI38V13/2RA while in the rest of the cell lines there was an increase in the level of C-circle DNA. Dot blot after hybridisation with a ^{32}P radiolabelled telomere probe. Signal intensities were measured using dot blot analyser from ImageJ software. Data was analysed using GraphPad Prism 5 software. A.U. Arbitrary Units. Replica of the assay on the same DNA showed the same results (Appendix 6 section A6.5).

5.8 Discussion

CtIP is a gene with a key role in HR. Since some genes involved in this mechanism are important participants in the ALT pathway this prompted the idea to investigate the effect of *CtIP* depletion in ALT cell lines. So far the only evidence linking *CtIP* with telomeres comes from studies in yeast showing its importance in C-strand processing and telomere elongation when short telomeres are detected (Bonetti D *et al.* 2009, Iglesias N *et al.* 2009). Here I have used immunofluorescence analysis to show that CtIP is often associated with telomeric DNA in ALT cells compared to telomerase+ cells. This association

could be at chromosome ends, with extrachromosomal DNA or both. At chromosome ends CtIP could be participating in 5' C strand processing to generate long 3' tails to promote strand invasion for telomere elongation through recombination or in repairing DSBs generated by stalled replication forks (Blackford AN *et al.* 2012). On the other hand APBs are possible sites of telomere extension in ALT cells, under this view CtIP could be participating in 3' tail generation. The known enhancing effect of CtIP on the nuclease activity of the MRN complex, which is indispensable for ALT, seems useful in these scenarios.

When colony formation efficiency was evaluated the shCtIPA did not seem to have an effect on WI38V13/2RA cells and a modest effect on SUSM1 cells, but the shCtIPB reduced significantly the number of colonies formed in both cell lines. The difference between the shRNAs could be related to an off target effect. However, the sequence of the shCtIPB was checked with a tool developed by the National Institute of Cancer (USA) and did not seem to align with any other transcripts (Ryan MC *et al.* 2008). In addition to the colony forming efficiency, the colony growth rate was monitored. Interestingly in both cell lines control clones grew at similar rates. Some of the shCtIP clones recovered from the SUSM1 cell line showed slow growth. For example clone shCtIPA1 took almost twice the time to reach PD20 compared to controls. However, qPCR analysis did not show convincing evidence of *CtIP* downregulation. On the other hand the only shCtIPA clone recovered from the WI38V13/2RA transfected cells also had slow growth and it showed 50% of *CtIP* downregulation. Nevertheless, the possibility that clones with higher levels of downregulation have not survived cannot be ruled out. In the same way the possibility that those clones with no evidence of downregulation at the time of analysis did have reduced expression of *CtIP* on stages cannot be discarded. It has been reported that *mCtIP*^{-/-} mouse embryos die at early stage of development and also 3T3 mouse cells die approximately one week after *CtIP* depletion with shRNA. Thus cell lethality seems to be caused by failure to progress into S phase and the inability to repair DSBs (Li S *et al.* 2000), (Chen PL *et al.* 2005), (Liu F *et al.* 2006).

As telomeres in ALT+ cells are maintained by a HR-like mechanism the generation of 3' tails long enough to promote strand invasion is required. Bonetti *et al*, showed a transient shortening of native telomeres in yeast *Sae2/Sgs1* double mutant which was reverted after approximately 75 cell divisions (Bonetti D *et al.* 2009). To determine if there were some effects at telomeres after *CtIP* downregulation different markers were evaluated. Both STELA and TRF were used to evaluate changes in telomere length in the clone shCtIPA1. The XpYp telomere was shown to be very short by STELA analysis in clone shCtIPA1. However the 12q and 17p telomeres were not different between the control and shCtIPA clones. This suggests that the short XpYp telomere in shCtIPA1 was not related to the *CtIP* reduction. In addition TRF analysis did not show changes in telomere length between the clones studied.

Telomere mutations were not detected because TVR maps from XpYp and 17p telomeres were not very informative. In addition it seems that the WI38V13/2RA shCtIPA1 has lost one sex chromosome since all the mapped STELA products had the same pattern of variant repeats. Furthermore, the 12q Δ allele seems to be lost in the shCtIPA1 clone. MS32 instability is a hallmark of ALT activity (Jeyapalan JN *et al.* 2005). In recent experiments was observed that BLM downregulation increased MS32 mutation rates (Mendez-Bermudez *et al*, unpublished data). However, the mutation rate in the MS32 minisatellite did not change upon 50% *CtIP* downregulation in the shCtIP clone.

The presence of extrachromosomal DNA circles containing telomeric sequences (t-circles) has been widely studied and identified. They have been identified in different organisms from yeast to mammalian cells. Some authors have proposed that they are conserved because they play an important role in telomere maintenance (Tomaska L *et al.* 2009). In human cells t-circles have been detected in ALT cells but they have also been found in telomerase+ cells after overexpression of the *hTERT*, associated with extensive telomere elongation. This also occurs in male germ cells where telomerase is active, suggesting that they are the result of a telomere trimming mechanism (Pickett HA *et al.* 2009), (Pickett HA *et al.* 2011). In 2009 Nabetani A *et al*, provided evidence of the existence of

different forms of extrachromosomal DNA in ALT+ cells. Using 2D gels they identified a structure different to the t-circles that they called ss-C. This was detected under non-denaturing conditions with a G-rich probe. (and migrated with single stranded and double stranded DNA markers). They proposed that this structure could be a C-strand circle with some regions of dsDNA originated either from processing of the chromosome end or t-circles. They also proposed that these circles could be substrates for telomere elongation via a rolling circle mechanism (Nabetani A *et al.* 2009). Further studies by Henson JD *et al.*, established that these circles were linked to ALT activity as they were detected only in ALT cells, they were dramatically reduced after depletion of key genes for ALT such as *RAD50* and they were found even in the absence of other ALT markers. The detection of these C-circles in blood from patients with ALT+ tumors may make them a good diagnostic tool (Henson JD *et al.* 2009). If C-circles are a reliable indicator of ALT activity then disturbance of the pathway should impact their detection. In this study the C-circle assay on the shCtIP1 clone with 50% downregulation of CtIP expression, did not detect a change in the C-circle abundance. The possibility that a change occurred earlier during clonal expansion and reverted due to redundant activity of other enzymes could not be ruled out. To investigate if the effect was immediate, cells were treated with siRNA, and a reduction in the abundance of C-circles observed after 48hrs of siCtIP treatment in the SUSM1 and W138V13/2RA. And this seemed to support an effect in C-circle formation. It is worth mentioning that reduced C-circle levels in the W138V13/2RA cell line had been reported previously (Henson JD *et al.* 2009). However when I repeated the experiment including two more ALT cell lines the results were completely opposite showing an increase in C-circles. This controversy support opposite roles for CtIP. One is that it has a key role in the ALT pathway and its depletion cause problems in ALT activity reflected by the reduction of C-circles. The other is that *CtIP* depletion promotes ALT activity reflected by the increase in C-circles, a problem during DNA extraction seems unlikely because that would imply degradation of circles exclusively in the siCtIP treated cells in the three ALT+ cell lines. One possibility is that the C-circle assay was not reliable in the laboratory and the results need to be corroborated by a different method. An alternative approach could

be the use of 2D gels as described by Nabetani *et al* (Nabetani A *et al.* 2009). A disadvantage of this approach is that 2D gels demand high amounts of DNA. Unfortunately it was not possible to optimize a method to analyse the single stranded telomeric 3'overhang, we were unable to clarify if CtIP is playing a role in the ALT mechanism.

In summary immunofluorescence demonstrated that CtIP associates with telomeric DNA in ALT cells but not in telomerase+ cells, suggesting that it plays a role in telomere maintenance. The lack of changes in the ALT markers studied could be related to the existence of more than one pathway as suggested by Bonneti *et al.* Downregulation of *CtIP* in conjunction with *BLM* or *WRN* could be helpful to investigate if these pathways observed in yeast are conserved in humans.

5.9 Future work

To date no one has reported a possible participation of CtIP in the ALT mechanism in human cells. However its role in HR repair and its interaction with the MRN complex makes it an attractive candidate. The association of CtIP with telomeric DNA in ALT+ but not in telomerase+ cell lines suggests that CtIP is involved in the ALT mechanism. Therefore the absence of a detectable effect on various markers of ALT is surprising and requires further investigation.

- Does CtIP localize to APBs? The exact role of APBs in the ALT mechanism is not fully understood yet, but the MRN complex has been found in APBs and because CtIP enhances its nuclease activity it would be important to investigate if CtIP is present in APBs. If CtIP is found in APBs it would be interesting to investigate if it is recruited by the MRN complex.
- Association of CtIP with telomeric DNA during the cell cycle. The association of CtIP with telomeric DNA was observed in asynchronous cultures. ALT activity is thought to occur on late S/G2 phase (Cesare AJ *et al.* 2010). Identifying whether CtIP could

be preferentially associated with telomeric DNA at a specific phase would be important to understand if it is playing a role in ALT.

- Incidence of telomere-telomere fusions, telomere free ends and META-TIFs. Based on experiments in yeast, Sae2, the CtIP yeast homolog, is important for the formation of 3' single stranded tails by degradation of the C strand (Bonetti D *et al.* 2009). The 3' single stranded tails are necessary for strand invasion and BIR progression (Llorente B *et al.* 2008). If the ALT mechanism resembles the BIR pathway depletion of CtIP could compromise elongation of short telomeres and promote telomere fusions, loss of the telomeric signal or changes in damage signalling.
- Abundance of POT1 after CtIP depletion. RPA is recruited to DSB sites to coat the ssDNA generated for initiate the repair process. The recruitment of RPA is significantly impaired after depletion of CtIP or expression of a mutant for of the protein (Huertas P *et al.* 2009). At telomeres POT1 is bound to the telomeric ssDNA, so it is possible that reduced levels of CtIP affect the presence of POT1 at telomeres. As reduction in the amount of POT1 should mean less ssDNA at telomeres. This effect could also be analysed by measuring the amount and/or length of the 3' single stranded telomeric overhang DNA after CtIP depletion. It is necessary to implement a method to evaluate changes in the amount and/or length of the 3' telomeric overhang.
- Conditional knockdown of *CtIP*. In this study just one clone from the WI38V13/2RA cell line ~50% of *CtIP* downregulation was generated. It is possible that higher levels of downregulation compromise cell viability. The use of a vector with controlled expression of the shRNA may help to study the effect of higher levels of *CtIP* depletion. In addition, the transfection efficiency should be improved in order to recover more clones and generate more robust data.
- Simultaneous depletion of CtIP and BLM. Based on experiments on yeast, Bonneti *et al.*, have proposed that at least two pathways control the C-strand resection at telomeres. One is governed by Sgs1 (BLM/WRN yeast homolog) and the other one

driven by Sae2 (CtIP yeast homolog). The Sgs1 Sae2 double mutant show severe impairment of C-strand processing. (Bonetti D *et al.* 2009). The generation of an ALT+ cell, with simultaneous depletion of CtIP and BLM using siRNA may help to investigate if these pathways are conserved in higher eukaryotes.

CHAPTER 6 – CHROMOSOME INTEGRATED HUMAN HERPESVIRUS 6 (CIHHV-6) MECHANISM AND CONSEQUENCES.

6.1 Background

Proof of chromosomally integrated HHV (CIHHV-6) into human telomeres was provided by Arbuckle *et al* in 2010. The authors also showed that the virus integrates with the DR_L towards the telomere and the DR_R towards the centromere (Chapter 1 Figure 1.10). Interestingly they did not find non-integrated episomes in infected cells and they claimed that the virus could be reactivated after treating the cells with the histone deacetylase (HDAC) inhibitor trichostatin A (Arbuckle JH *et al.* 2010). This evidence supported the idea that the latent state of HHV-6 is reached through chromosomal integration. However, the debate about HHV-6 latency is still on, and several authors have pointed out that stronger evidence is needed. In a recent review article a worldwide frequency of CIHHV-6 close to 1%, with a 1:3 ratio of HHV-6A:HHV-6B was reported (Morissette G *et al.* 2010).

After the confirmation that HHV-6 can integrate into telomeres the mechanism of integration has begun to be investigated. Among others, two features of the HHV-6 genome have been the focus of attention, one is the T2 region (composed of perfect telomeric repeats) and the second is the gene U94. There are a few examples of herpesviruses with the capacity to integrate their genomes into human chromosomes. Marek's disease virus (MDV) is one example of a herpesvirus able to integrate into host telomeres and it also contains telomeric repeats. MDV is considered an oncogenic virus and its integration contributes to the development of chicken T-cell lymphomas. Its genome consists of one unique long (U_L) and one unique short (U_S) regions and, each is flanked by internal and external repeats. The external repeats contain telomeric repeats (TMR) and recently published data have shown that, following mutation or deletion of the TMR, the virus integrates in multiple intra-chromosomal regions instead of telomeres. And also it was observed that mutation or deletion of the TMR caused impairment in virus reactivation and tumor development in chickens (Kaufer BB *et al.* 2011). Even though the

integration mechanism is not fully understood, the data from MDV support the idea that telomeric integration is mediated by the viral TMR. However, the presence of telomeric repeats alone might not be enough for virus integration. It is known that human herpes virus 7 (HHV-7) has a similar genome organization to HHV-6. The HHV-7 genome also consists of a unique region flanked by direct repeats, which contain T1 and T2 elements (Black JB *et al.* 1999). In contrast to HHV-6, to date, there is no evidence of HHV-7 chromosomal integration. The arrangement of the conserved blocks of genes between HHV-6 and HHV-7 is identical, and the amino acid sequence identity within the core genes is between 41% and 75%, despite these similarities, remarkably neither HHV-7 or other human herpesviruses possess a homologue of the U94 gene which encodes a homologue of the human adeno-associated virus type 2 (AAV-2) Rep78/68 gene (Caselli E *et al.* 2007). Rep 78/68 displays at least three activities that may be related to virus integration: DNA binding, endonuclease and helicase-ATPase. U94 and Rep 78/68 show 24% amino acid identity, and partial characterization of U94 indicates that it has DNA binding capacity. The helicase and endonuclease activities of U94 have not been found, but key amino acids for these functions are conserved between U94 and Rep78/68 (Morissette G *et al.* 2010). In addition Thompson *et al.*, showed that U94 can complement replication of a Rep-deficient AAV-2, and they proposed that HHV-6 might have acquired a copy of the Rep gene from AAV-2 during co infection *in vivo* (Thomson BJ *et al.* 1994). Current evidence suggests that chromosomal integration of HHV-6 could occur by homologous recombination between the T2 element and human telomere repeats (Chapter 1 Figure 1.11). This model involves the participation of the shelterin complex and U94, but more information is needed to unravel the mechanism underlying the integration of HHV-6.

Equally important are the possible clinical implications of CIHHV-6. What advantages or disadvantages have the individuals born with an extra 160 kb of a foreign genome in each cell of their bodies? It is clear that HHV-6 shows an unusual behaviour for a human herpesvirus, but whether it is necessary or not to develop a routine test to screen all newborns to detect CIHHV-6 (in order to be aware of possible risks) is yet to be answered.

Basic and clinical research is being carried out all around the world to understand the causes and consequences of CIHHV-6.

6.2 Aim

To investigate the frequency of CIHHV-6 in worldwide populations and the mechanism of integration

6.3 CIHHV-6 Population screening

In this study a survey was carried out to investigate CIHHV-6 frequency in several populations. The survey was done using a PCR approach. Primers designed to amplify the U18 region (U18F and U18R see appendix 1) from either HHV6-A or B subtypes were used to screen genomic DNA samples. Positives samples were verified by amplifying a specific region that distinguishes the A and B subtypes (DR5 for HHV-6A and DR6 for HHV-6B primers in Appendix 1). DNA Samples from the HGDP-CEPH were derived from blood DNA, while the NEOCODEX Spanish samples composed whole genome amplification DNA. A total of 2645 samples were screened and the overall frequency of CIHHV-6 was 0.86% (23 positives), with a ratio A:B of 1:2.2 (Table 6.1). This frequency was close to the 1% frequency with an A:B ratio of 1:3 reported previously (Morissette G *et al.* 2010). To rule out the possibility that PCR screen detected an episomal form of the virus single telomere length analysis (STELA) was carried out on positives samples. A subtelomeric primer (DR1R see Table 2.1), within the DR1 region of HHV6 was used (Royle NJ, *et al*) in the STELA reactions with the Teltail and Telorette2 primers. In this way the STELA reactions were made specific for integrated virus genomes (Figure 6.1). This confirmed that the 19 HHV-6 positive samples in the HGDP-panel are all from CIHHV-6 donors. Unfortunately DNA from whole genome amplification is not suitable for STELA and so integration could not be confirmed in the four Spanish samples that were positive for HHV-6 sequences.

Interestingly no HHV-6 positive samples were detected in African, American and Oceanian populations, although the number of samples from each population is low so the absence of HHV-6 could just be a sampling error or, reflect lower frequency of CIHHV-6 in these populations. On the contrary the frequencies in Sardinia and Orkney, 7.1% and 4.7% respectively, were remarkably high (Table 6.1). Again this may be due to sampling error but it is interesting to note that both populations are from islands. This raises the possibility of a Founder effect within a relatively isolated population or possibly the presence of particular strains of the virus which integrate more efficiently. It is also interesting that in the Spanish population the ratio A:B is inverted being 3:1. Again this raises the possibility that some populations are more susceptible to HHV-6 integration or that some strains of the virus integrate more efficiently. More samples from Orkney and other populations have been screened (Yan Huang, Nicola Royle personal communication) to get a better picture of the frequency of CIHHV-6. In addition some regions of the virus genome are being sequenced in order to generate phylogenetic trees of the integrated viruses.

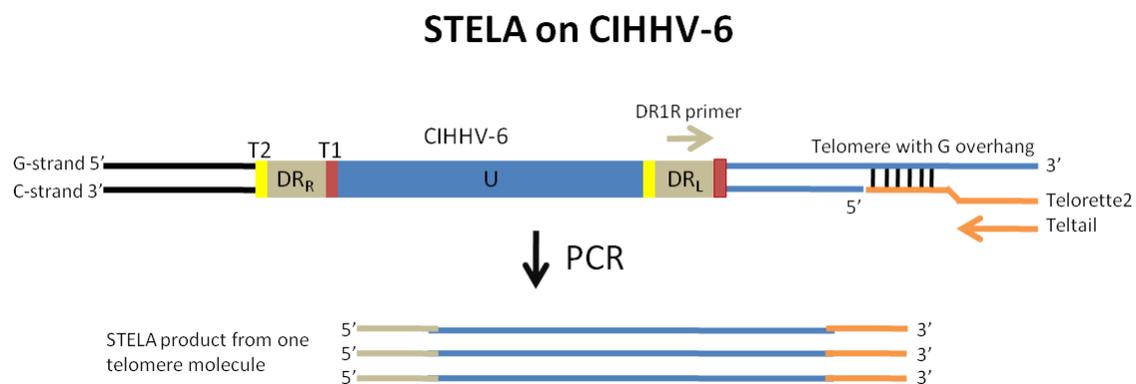


Figure 6.1 STELA on HHV-6 positive samples was used confirm the presence of an integrated copy of HHV-6. The DR1R primer anneals to the DR1 region of both HHV-6 A and B. It can be substituted as the subtelomeric primer in STELA and works in combination with Teltail and Telorette2 to generate telomere amplicons. The STELA products generated contain the DR1 and the T1 regions at one end and the human telomeric repeats at the other end. T1 and T2 are the regions with imperfect and perfect telomeric repeats respectively. U corresponds to the HHV-6 Unique region.

Table 6.1 CIHHV-6 frequencies in the HDGP-CEPH panel and Spanish samples. Samples from the British Isles were kindly provided by Dr. Bruce Winney and Professor Sir Walter Bodmer, University of Oxford. The Spanish samples were obtained from NEOCODEX by Professor Rhona Borts (Department of Genetics, University of Leicester).

Region	Samples	Positives	HHV6A	HHV6B
Africa				
Sub-Saharan Africa	179	0	0	0
North Africa	30	0	0	0
Europe				
North/West Europe (CEPH)	148	1.3% (2)	0	2
British	544	1.2% (7)	1	6
Sardinia	28	7.1% (2)	0	2
Orkney	42	4.7% (2)	0	2
Spain	671	0.59% (4)	3	1
Middle East (Israel)	163	1.2% (2)	2	0
Central/South Asia				
Pakistan	192	0.5% (1)	0	1
China	10	0	0	0
East Asia				
China	222	0.9% (2)	0	2
Japan	76	1.3% (1)	1	0
Siberia	24	0	0	0
Oceania				
Bougainville	17	0	0	0
New Guinea	11	0	0	0
America				
Brazil	45	0	0	0
Colombia	13	0	0	0
Mexico	50	0	0	0
Total	2645	0.86% (23)	7	16

CIHHV-6 can be inherited in a Mendelian fashion in families but so far no studies have been conducted to analyse the frequency of CIHHV-6 in the germline. Here a panel of 92 sperm DNA samples from British donors was screened for the presence of HHV-6. The survey revealed a frequency of 5.4% (5 positives). 1 HHV-6A and 3 HHV-6B, unfortunately there was no more DNA available from the fifth positive to determine which variant it contained (Table 6.2). It was not possible to determine which chromosome was harbouring the CIHHV-6 because only DNA was available so it was not possible to do FISH.

The difference between the frequencies observed in the germline DNA (5.4%) and in somatic DNA from the British Isles samples (1.2%) is statistically significant ($p=0.01$ Fisher's exact test). Some studies have detected the presence of HHV-6 in semen, confirming that HHV-6 is able to cross the hematotesticular barrier and reach the germ cell compartment (Michou V *et al.* 2011). This raises the possibility of direct integration of HHV-6 into germ cells. One possible explanation for this differences between somatic and germline frequencies are a sampling error because the low number of sperm samples available. However, it also raises the possibility that HHV-6 would confer some disadvantage to gametes or embryos harbouring CIHHV-6, reducing the likelihood of vertical transmission of the CIHHV-6. To investigate this last possibility, it would be necessary to study a large number of families with CIHHV-6 and see if it is transmitted to the offspring according to Mendel's segregation law, and also to investigate whether CIHHV-6 can cause pregnancy loss.

Table 6.2 Frequency of CIHHV-6 in a panel of 92 DNA samples from sperm from British donors.

DNA from sperm				
Individuals	Positives	Frequency	HHV6A	HHV6B
92	5 ^x	5.4%	1	3

^xDNA from the fifth donor was not available to determine the variant.

6.4 Telomere length analysis on chromosomes with CIHHV-6 from somatic and germline DNA

6.4.1 Telomere length analysis on chromosomes with CIHHV-6 from somatic DNA

It is known that telomere length is one of the key factors in maintaining telomere function and genomic stability. So far no studies have investigated whether the insertion of the 160Kb HHV-6 genome into a telomere affects telomere length or its maintenance by

telomerase. To address these questions telomere length analysis was carried out using STELA on DNA from somatic cells and on sperm DNA.

Blood DNA from three siblings was kindly provided by Professor Martin Dyer. A 73 year old woman diagnosed with Primary Effusion Lymphoma (PEL) (ID sample 1500) and her two clinically asymptomatic brothers (unknown ages, ID samples 1499 and 1501). CIHHV-6A chromosome 19q had been identified in all three donors but no data or samples were available from other relatives (Clark DA *et al.* unpublished data). Telomere length on the chromosome carrying the CIHHV-6 (referred as DR1R telomere hereafter) was compared with three other telomeres XpYp, 12q and 17p using STELA. In the three siblings the mean length of the DR1R telomere was the shortest among the telomeres studied (Figure 6.2, 6.3 and Table 6.3). The trend for the DR1R telomere to be the shortest among the studied telomeres agrees with previous results obtained from telomere length analysis performed in a lymphoblastoid cell line harbouring CIHHV-6A (10q telomere) in which the DR1R telomere was the shortest (Gunjan Mukesh Wig, unpublished data). Statistical analysis of median telomere lengths, using one way ANOVA with Newman-Keuls multiple comparison test (Table 6.3) showed that the difference between the DR1R and the 12q and XpYp telomeres is statistically significant in the three siblings. On the other hand the difference between the DR1R and 17p telomere was not statistically significant only in donor 1500 ($p > 0.05$). This observation was not completely unexpected because it has been previously reports that the 17p telomere is often shorter than the total telomere median (Martens UM *et al.* 1998). Currently more lymphoblastoid cell lines harbouring CIHHV-6A are being studied to see if this observation can be corroborated.

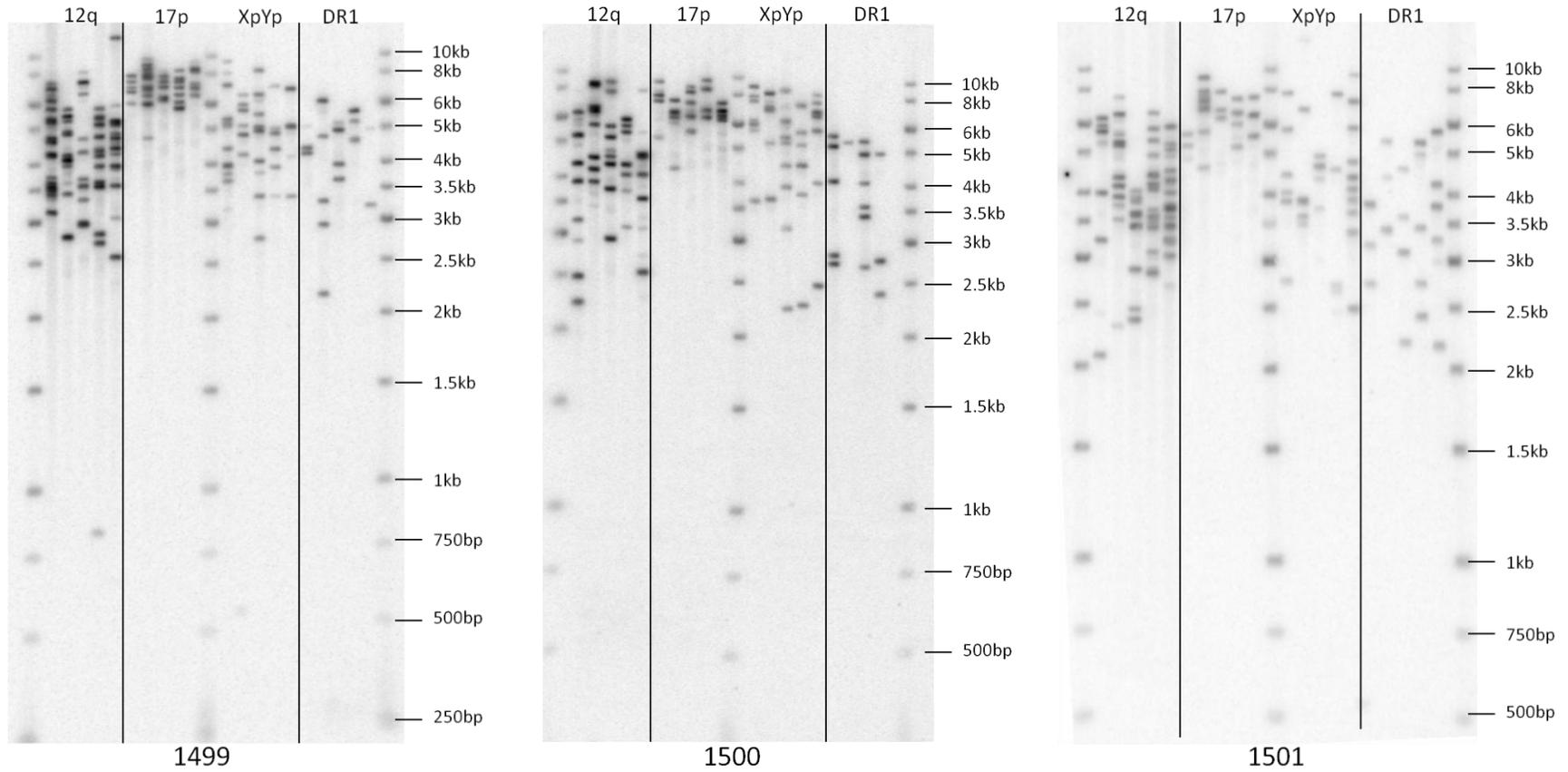


Figure 6.2 Examples of STELA Southern blots from siblings 1499, 1500 and 1501 used for telomere length analysis. Genomic DNA from each sibling was digested with *EcoRI* and 250 μ g of digested DNA were used per reaction. STELA was conducted using specific subtelomeric primers for telomeres DR1R, 12q, 17p and XpYp with the primers Teltail and Telortte2. The products were detected following agarose gel size fractionation and Southern blot hybridisation to a 32 P radiolabelled telomere repeat probe. Sizes of the STELA products were calculated using the ImageQuant software (Amersham). Additional gels were carried out to complete the analysis (appendix 7 section A7.1)

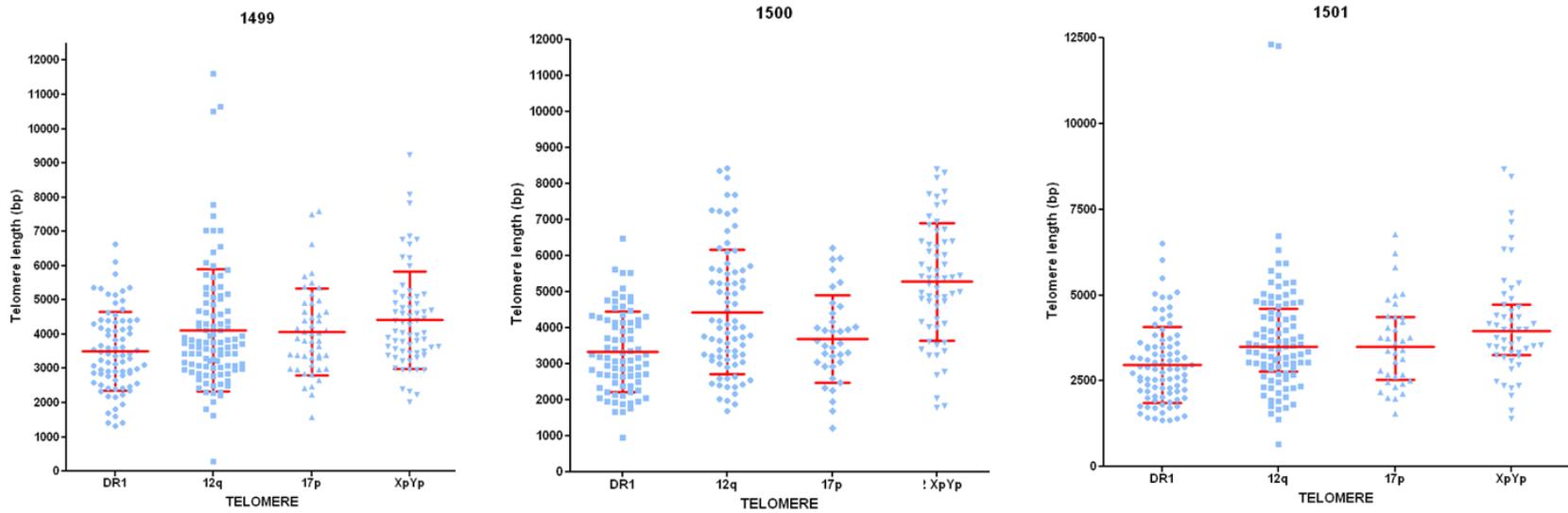


Figure 6.3 Telomere length analysis from blood DNA from three siblings carrying CIHHV6-A at 19q, means with standard deviations are shown. DR1R represents the telomere harbouring the HHV-6A integration. Analysis was carried out after subtraction of the flanking DNA in each STELA product 415bp for XpYp, 3078bp for 17p, 732bp for DR1R and 544bp for 12q. Data were analysed using GraphPad Prism 5 software.

Table 6.3 Comparison of mean telomere length between the CIHHV-6A telomere (DR1R) and 12q, 17p and XpYp telomeres from the three siblings carrying the virus at 19q. One way ANOVA with Newman-Keuls multiple comparison test was used. In each row the value on the left corresponds to the DR1R mean of the respective donor. The difference between the DR1R and the 12q and XpYp telomeres was statistically significant in the three siblings. The difference between the DR1R and 17p telomeres was not significant only in the donor 1500.

Mean comparison of DR1R telomere			
	Vs 12q	Vs 17p	Vs XpYp
1499	3497 Vs 4125 p = 0.05	3497 Vs 4069 p = 0.05	3467 Vs 4411 p = 0.01
1500	3333 Vs 4426 p = 0.0001	3333 Vs 3693 p > 0.05	3333 Vs 5274 p = 0.0001
1501	2964 Vs 3747 p = 0.001	2964 Vs 3582 p = 0.05	2964 Vs 4160 p = 0.0001

One important issue when calculating telomere length from STELA products is to adjust the length of each molecule by subtracting the length of the non-telomeric DNA between the location of the primer and the start of the telomere marked by the first telomeric repeat (i.e. distance in bp). For the 12q, 17p and XpYp the length of the sequence subtracted is 544, 3078 and 415bp respectively. For the HHV-6 associated telomere 732bp where subtracted since according to the consensus sequence of HHV-6A (NCBI Ref Seq: NC_000898.1) this is the distance from the DR1R1 primer to the end of the DR_L. However recent experiments have showed that the T1 region is expanded in both variants, more importantly in the B variant reaching up to 8kb (Royle NJ unpublished data). This T1 expansion makes it difficult to exactly determine the distance from DR1R primer to the end of the viral genome.

6.4.2 Telomere length analysis on chromosomes with CIHHV-6 in the germline

To investigate whether telomerase is able to maintain and extend CIHHV-6 telomeres adequately in the germline, telomere length analysis was conducted at 12q, 17p, XpYp and HHV-6 in the DNA extracted from the four sperm donors identified above. The analysis showed that HHV-6 associated telomeres were long, as expected in the germline. In contrast to the analysis of the blood DNA from the three siblings (above) the HHV-6 associated telomere was not the shortest in the germline. Furthermore in two donors, donor 32 and donor 56, the median of the HHV-6 associated telomere was the longest (Figures 6.4 and 6.5). To what extent this is an effect of the expansion of the T1 region is difficult to determine since data obtained in Dr. Royle's laboratory has shown that the length of the T1 region varies among the same subtype (Dr. Royle *et al*, unpublished data). However this investigation has shown that CIHHV-6 does not seem to affect telomerase function in the germline.

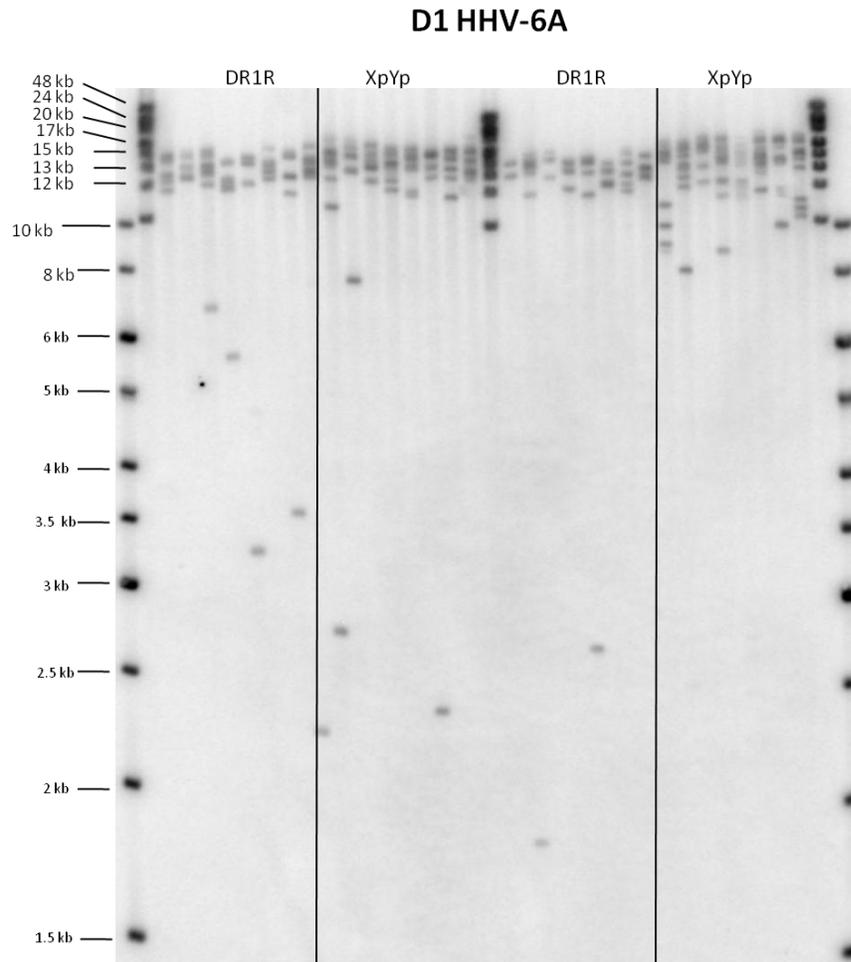


Fig 6.4 STELA Southern blot analysis in sperm DNA from donor 1 (D1). Example of STELA PCR using the HHV-6 (DR1R) and the XpYp telomere are displayed. Genomic DNA was digested with *EcoRI* and 300 pg of digested DNA were used per reaction (per track). The products were detected following agarose gel size fractionation and Southern blot hybridisation to a ^{32}P radiolabelled telomere repeat probe. Sizes of the STELA products were calculated using the ImageQuant software (Amersham). All the Southern blots for this analysis are shown in Appendix 7 section A7.2.

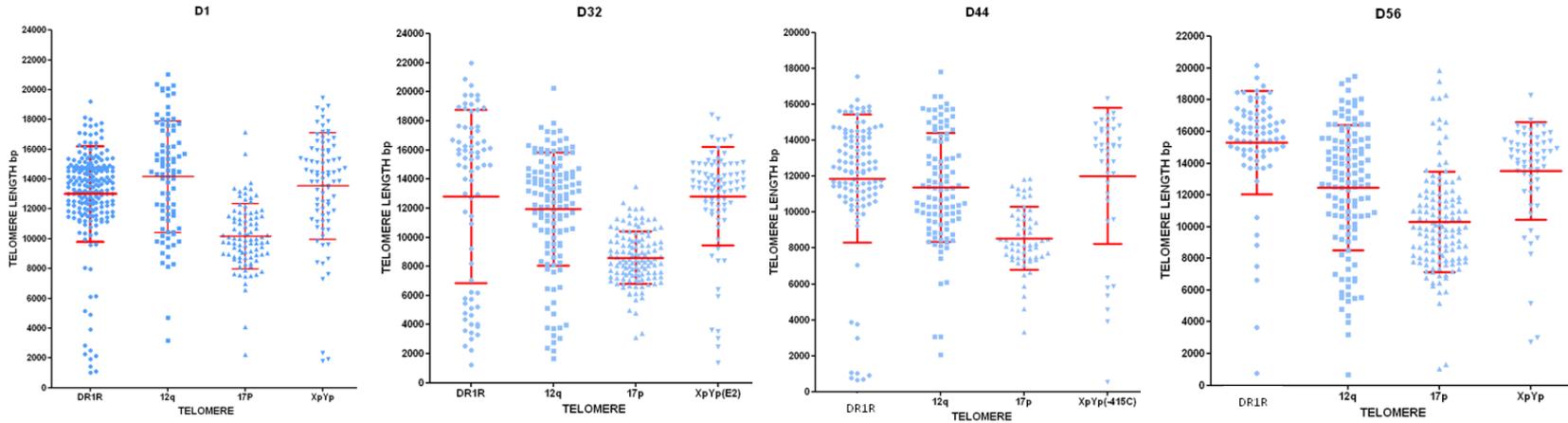


Figure 6.5 Telomere length analysis on DNA from germline after adjustment for flanking DNA in each STELA product 415bp for XpYp, 3078bp for 17p, 732bp for HHV-6 (DR1R) and 544bp for 12q. Data were analysed using GraphPad Prims 5 software. Means and standad deviations are displayed. Telomeres harbouring CIHHV-6 are not the shortest which means that they are being extended by telomerase in the germline.

6.5 Analysis of short telomeres on chromosomes with CIHHV-6 from somatic and germline DNA

In 2006 Baird DM *et al*, used STELA to investigate XpYp and 17p telomeres in the germline. They showed that the frequency of short telomeres (defined as those telomeres with length below the mean-2.33SD) in male germline was 3.6%. The percentage expected was 1% because according to a normal distribution this should be the percentage of molecules 2.33 standard deviations (SD) below the mean (Baird DM *et al*. 2006). These short telomeres represent truncated molecules that were either not properly extended by telomerase in primary spermatocytes or could be the result of telomere rapid deletions. It was proposed that if present in a zygote they may have negative consequences by limiting the proliferative potential of zygotes receiving these chromosomes. Some studies have proposed that various factors can influence telomere length at specific chromosomes, with some of them acting *in cis*. For example the abundance of the shelterin members TRF1 and the POT1 seem to regulate telomere elongation by telomerase (Smogorzewska A *et al*. 2004). In addition, different authors have reported that telomere length have a heritability index between 36-90%. Even though this is a wide range it is clear that genetic factors are involved in determining telomere length (Monaghan P 2010). Linkage studies seem to point to a locus at chromosome 12p influencing telomere length, the candidate gene within this region was the *DDX11* helicase (Vasa-Nicotera M *et al*. 2005). On the other hand epigenetic factors also can be involved in the control of telomere extension by telomerase. It has been demonstrated that patients with ICF syndrome (caused by mutations on the DNA methyltransferase-3B) the lower levels of methylation are associated with a shorter telomeres compared to age matched healthy controls (Yehezkel S *et al*. 2008). In order to investigate whether the presence of a CIHHV-6 affected the frequency of short telomeres in the germline I decided to analyse this in four CIHHV-6 positive sperm donors. In the study published by Baird DM, *et al*, they treated their samples as having a normal distribution and so they considered those molecules 2.33SD below the mean as the unexpected short molecules. To check if the telomeres showed

normal distribution, data sets (telomere lengths) of each telomere were plotted in a histogram, to see if they resemble normal distribution (Appendix 7 section A7.3). Most of the telomeres resembled normal distribution. The number of molecules below 2.33SD from the mean was compared between the HHV-6 associated telomere and the 12q, 17p and XpYp telomeres within each donor (Figure 6.6, Table 6.4 and 6.5). The histogram of the HHV-6 associated telomere from donor 32 showed two populations of molecules and this can be also seen in the telomere length analysis (Figure 6.5). I considered this telomere as especially long, the population of longest molecules (>10kb) could represent those telomeres being extended by telomerase, while the shortest molecules could be the result of a trimming mechanism to regulate telomere length in the germline, with no relation to the HHV-6 integration. The group of molecules above 10kb seems to have normal distribution, therefore, the mean telomere length in the HHV-6 associated telomere from donor 32 was calculated using this population of molecules. The analysis revealed that besides donor 32 there is no statistically significant difference in the proportion of telomeres below 2.33SD from the mean among the analysed telomeres. The difference observed in donor 32 could be related to the inherent nature of the telomere for being especially long and not to the HHV-6 integration and should be taken with caution.

In addition to the analysis of the molecules below 2.33SD from the mean, I decided to analyse the telomeres at a point where telomeres become dysfunctional. According to different authors telomeres become dysfunctional when the length is $\leq 1.5\text{kb}$ leading to uncapping and consequent genomic instability (Palm W *et al.* 2008, Counter CM *et al.* 1992, Mathieu N *et al.* 2004). Therefore I compared the proportion of molecules $\leq 1.5\text{kb}$ among the studied telomeres in the four sperm donors. STELA was used for the analysis of short telomeres in the four sperm donors. The STELA products and gels used for this analysis were different from those used to measure telomere length. The reason to make new gels was because in the gels used to measure telomere length the molecules below the 1.5kb DNA marker were lost (Figure 6.4). For the analysis of short telomeres the

electrophoresis was carried out in a way the 250bp band of the DNA stayed on the gel.

Table 6.4 Analysis of telomeres below 2.33SD from the mean telomere length. The number of telomeres 2.33SD below the mean telomere length in the HHV-6 associated telomere (DR1R) was compared against the number in the 12q, 17p and XpYp telomeres in the four sperm donors. Total number of telomeres from each telomere and donor are shown in the last column. In parenthesis is the respective percentage to the number of telomeres above and below 2.33SD from the mean telomere length.

Donor	Telomere	Mean telomere length	Telomeres >2.33SD from the mean (%)	Telomeres <2.33SD from the mean (%)	Total telomeres (100%)
D1	DR1R	12993	194 (94.6%)	11 (5.4%)	205
	12q	14162	81 (97.5%)	2 (2.5%)	83
	17p	10155	89 (97.8%)	2 (2.2%)	91
	XpYp	13532	79 (96.3%)	3 (3.7%)	82
D32	DR1R	16629	49 (67.2%)	24 (32.8%)	73
	12q	11940	126 (96.9%)	4 (3.1%)	130
	17p	8591	123 (98.4%)	2 (1.6%)	125
	XpYp	12820	78 (93.9%)	5 (6.1%)	83
D44	DR1R	11880	102 (93.5%)	7 (6.5%)	109
	12q	11390	99 (97%)	3 (3%)	102
	17p	8547	57 (98.2%)	1 (1.8%)	58
	XpYp	12020	40 (97.5%)	4 (2.5%)	41
D56	DR1R	15302	73 (97.3%)	2 (2.7%)	75
	12q	12474	121 (98.3%)	2 (1.7%)	123
	17p	10291	125 (98.4%)	2 (1.6%)	127
	XpYp	13520	57 (95%)	3 (5%)	60

Table 6.5 Fisher’s exact test results. The number of telomeres below 2.33SD from the mean telomere length in the HHV-6 associated telomere (DR1R) were compared with the 12q, 17p and XpYp telomeres in the 4 sperm donors using Fisher’s exact test. $p \leq 0.05$ values are considered as statistically significant difference. Donor 32 is the only one showing significant differences.

Fisher’s exact test			
	DR1R Vs 12q	DR1R Vs 17p	DR1R Vs XpYp
D1	$p=0.24$	$p=0.16$	$p=0.56$
D32	$p=0.0001$	$p=0.0001$	$p=0.0001$
D44	$p=0.33$	$p=0.26$	$p=0.44$
D56	$p=0.055$	$p=0.19$	$p=1$

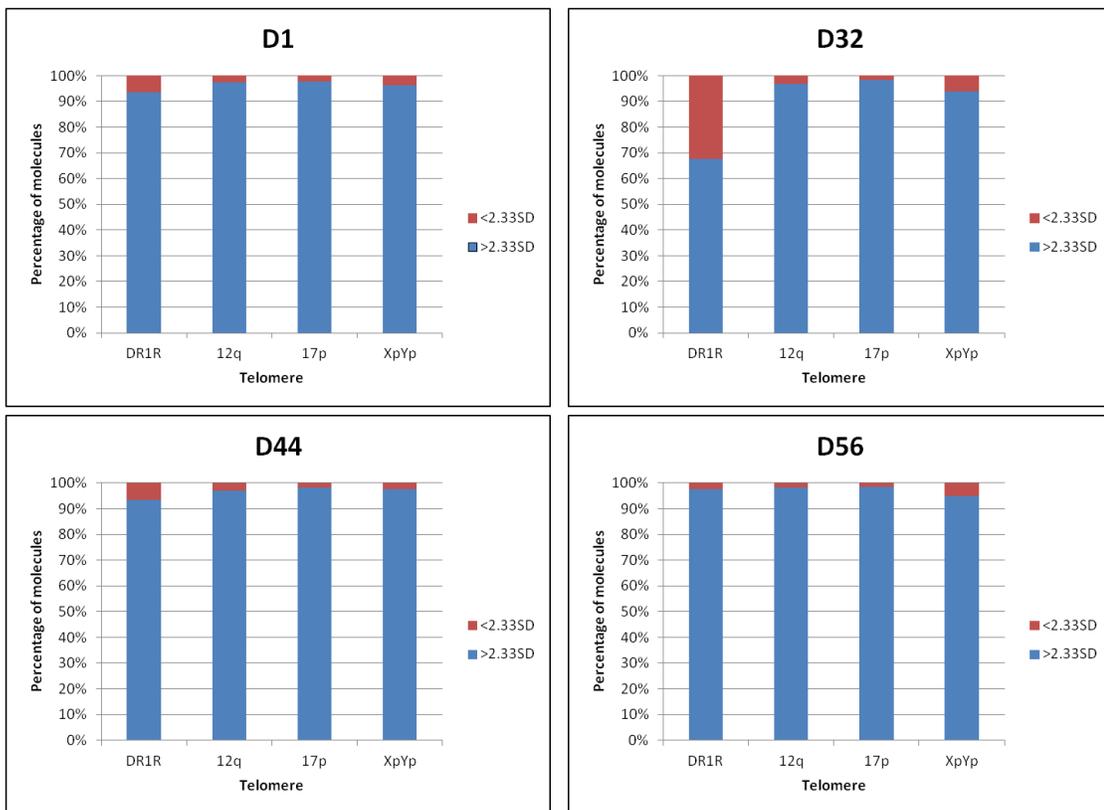


Figure 6.6 Analysis of telomeres below 2.33SD from the mean. The percentage of telomeres 2.33SD below the mean telomere length (in red) in the HHV-6 associated telomere (DR1R) was compared against the 12q, 17p and XpYp telomeres in the four sperm donors. In blue is the percentage of telomeres above 2.33SD the mean telomere length.

With this strategy even the shortest telomeres were retained (Figure 6.7). The number of molecules $\leq 1.5\text{kb}$ (after adjustment for flanking DNA) were counted in the HHV-6 associated telomere and compared with 12q, 17p and XpYp telomeres. Molecules $\leq 1.5\text{kb}$ were found in almost all the telomeres studied in all the subjects, with the exception of 17p in donor 1 and 56 and XpYp in donor 56. When the proportion of short molecules found in the HHV-6 associated telomere was compared with the other telomeres there was a statistically significant difference, showing an increase in the proportion of short telomeres compared to 12q and 17p telomeres in donor 1. In donor 32 there was a difference when the HHV-6 associated telomere was compared to the XpYp. In the donor 56 the difference was observed when the HHV-6 associated telomere was compared to the other three telomeres. However, in donor 44 there was no difference when the HHV-6 associated telomere was compared to the other three telomeres (Fisher's exact test) (Figure 6.8, Tables 6.6 and 6.7). The fact that the difference between the HHV-6 associated telomere and the rest of the 12q, 17p and XpYp telomeres is not always significant, and that in donor 44 no difference was observed, suggests that solely the presence of the virus is not enough to cause an increase in the number of short telomeres and that additional factors inherent to each chromosome and each individual participate in the truncation of telomeres in male germline.

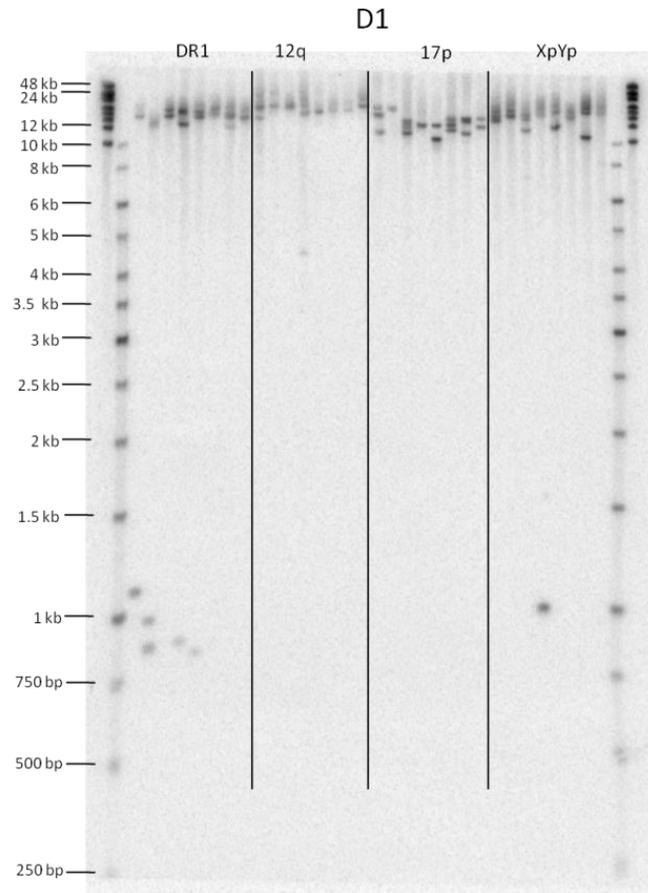


Fig 6.7 STELA Southern blot analysis in sperm DNA from donor 1 (D1) to measure short telomeres. Example of STELA PCR using the HHV-6 (DR1R), 12q, 17p and the XpYp telomeres primers with Teltail and Telorette2 are displayed. Genomic DNA was digested with *EcoRI* and 300 μ g of digested DNA were used per reaction (per track). The products were detected following agarose gel size fractionation by Southern blot hybridisation to a 32 P radiolabelled telomere repeat probe. Sizes of the STELA products were calculated using the ImageQuant software (Amersham). All the Southern blots used for this analysis are shown in Appendix 7 section A7.4.

Chromosome integrated human herpesvirus 6 (CIHHV-6) mechanism and consequences

Table 6.6 Analysis of short telomeres. STELA products were generated from the HHV-6 associated telomere (DR1R), 12q, 17p and XpYp telomeres from each donor to analyse the presence of short telomeres. Total number of telomeres from each telomere and donor are show in the last column. In parenthesis is the respective percentage to the number of telomeres above and below 1.5kb.

		Telomeres >1.5kb	Telomeres <1.5kb	Total
D1	DR1R	100 (87%)	14 (13%)	114
	12q	150 (94%)	9 (6%)	159
	17p	127 (100%)	0 (0%)	127
	XpYp	118 (93%)	8 (7%)	126
D32	DR1R	87 (97%)	2 (3%)	89
	12q	110 (96%)	5 (4%)	115
	17p	108 (95%)	5 (5%)	113
	XpYp	39 (88%)	5 (12%)	44
D44	DR1R	88 (93%)	6 (7%)	94
	12q	153 (92%)	13 (8%)	166
	17p	98 (96%)	4 (4%)	102
	XpYp	76 (93%)	5 (7%)	81
D56	DR1R	32 (78%)	9 (22%)	41
	12q	81 (93%)	6 (7%)	87
	17p	62 (100%)	0 (0%)	62
	XpYp	25 (100%)	0 (0%)	25

Table 6.7 Fisher's exact test results. The number of short telomeres in the HHV-6 associated telomere (DR1R) was compared with the number of short telomeres in the 12q, 17p and XpYp telomeres in the 4 sperm donors using Fisher's exact test. $p \leq 0.05$ values demonstrate there is a statistically significant difference.

Fisher's exact test			
	DR1R Vs 12q	DR1R Vs 17p	DR1R Vs XpYp
D1	p=0.04	p=0.0001	p=0.08
D32	p=0.05	p=0.05	p=0.03
D44	p=0.8	p=0.5	p=0.1
D56	p=0.01	p=0.0001	p=0.0001

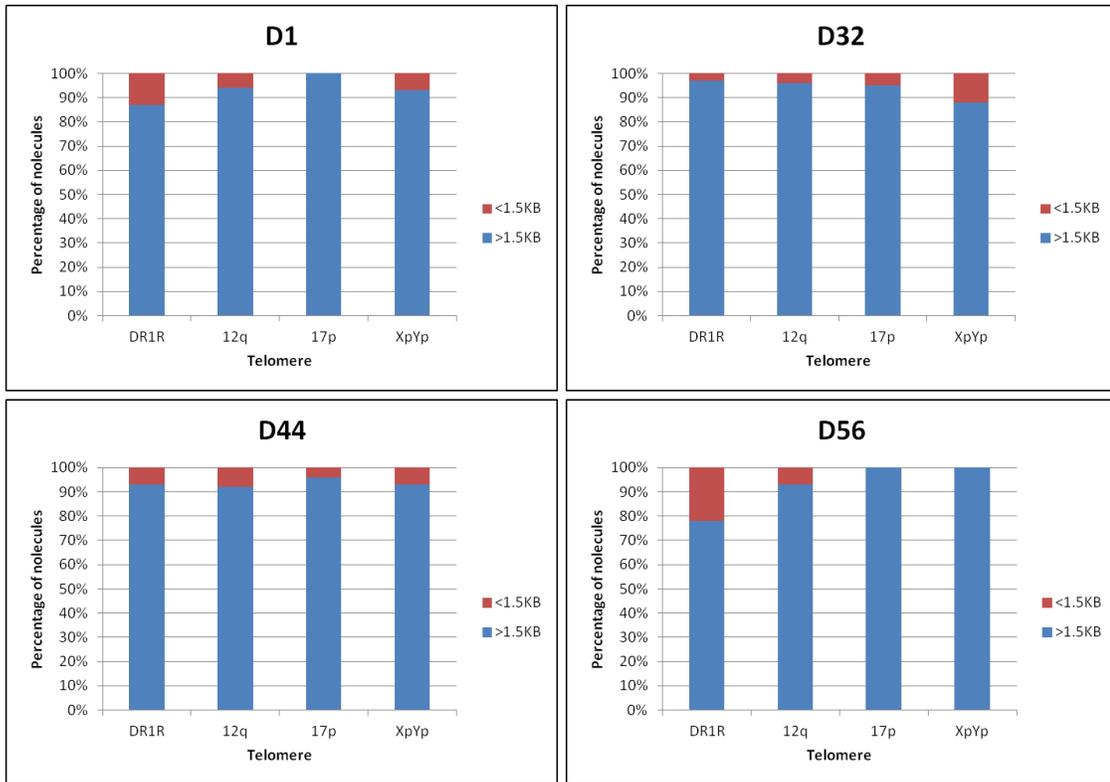


Figure 6.8 Analysis of short telomeres ($\leq 1.5\text{kb}$) found in DNA from the four sperm donors. STELA was used to amplify the 12q, 17p, XpYp and the HHV-6 associated telomere in the four sperm donors harbouring CIHHV-6. The percentage of molecules $\leq 1.5\text{kb}$ (in red) found in the HHV-6 associated telomere was compared against the 12q, 17p and XpYp telomeres within each donor. In blue are the percentage of telomeres $> 1.5\text{kb}$.

6.6 Sequence analysis of the proximal and distal breakpoints on chromosomes with CIHHV-6

6.6.1 Sequence analysis of the proximal breakpoint on chromosomes with CIHHV-6

The proposed mechanism for HHV-6 integration is through HR between the perfect telomeric repeats in the T2 region located at the end of the right DR (DR_R), and the human telomere. In this model the DR_R is brought next to the chromosome at the proximal junction (Figure 6.1). In 2010 Arbuckle JH *et al*, were able to PCR amplify the viral-human proximal junction in a family with CIHHV-6A at 17p by using a specific primer for 17p

subtelomere and one annealing at the end of the DR_R. The isolated junction sequence showed loss of the Pac2 element and just five TTAGGG repeats between the human and HHV-6 genome. In addition, they used inverse PCR (IPCR) to isolate a variety of viral-human DNA fragments in HEK-293 clones after infection with HHV-6 suggesting that the virus integrates efficiently into HEK-293s cell *in vitro* (Arbuckle JH *et al.* 2010).

In Dr. Royle's laboratory five lymphoblastoid (telomerase+ EBV transformed cell lines) with CIHHV-6 at known telomeres were available (Table 6.8). I used IPCR on KUK to isolate the HHV-6A-10q junction. A Southern blot of the IPCR products with a HHV-6 probe specific for the DR8 region showed a product of ~1.3kb at five annealing temperatures (58°C-65°C). Re-hybridisation with telomere probe showed this product also contained telomere like repeats (Figure 6.9A).

Table 6.8 Known location of the CIHHV-6 in the lymphoblastoid telomerase+ EBV transformed cell lines from five unrelated donors

Cell line	Inserted HHV-6 variant	Integration site as detected by FISH ^a
KUK	A	10q26.3
OL	B	9q34.3
AS	B	17p13.3
AMD	B	11p15.5
MP	B	17p13.3

^aFISH localisation described in (Nacheva EP *et al.* 2008), (Clark DA *et al.* 2006).

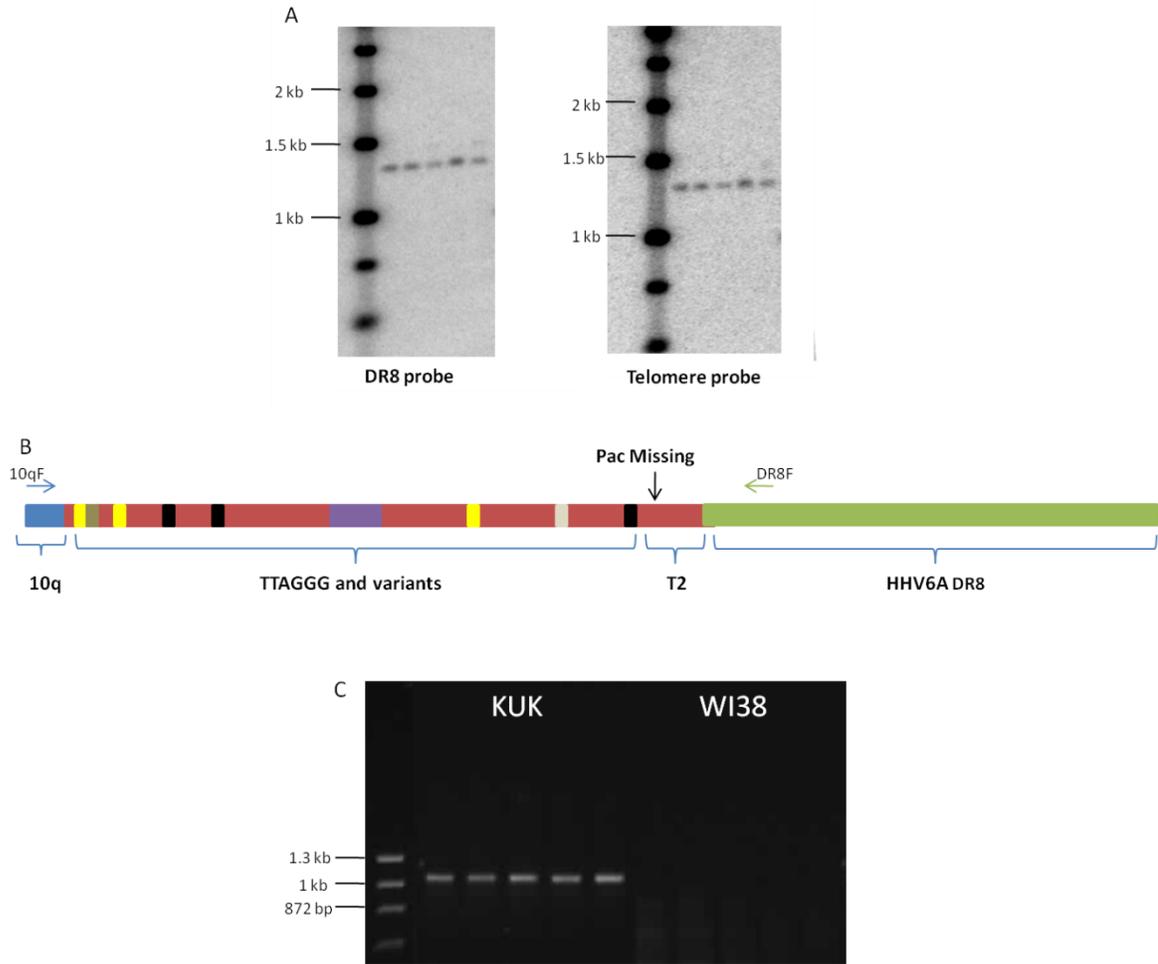


Figure 6.9. Internal junction of HHV-6A and human chromosome 10q was obtained by IPCR from KUK cell line. A) Hybridization of the IPCR products obtained at five different annealing temperatures (58°C to 65°C) with DR8 probe or telomere probe. B) Schematic representation of the sequence of the IPCR product (not to scale), the 10q sequence (blue) was found at one end, followed by a region of TTAGGG repeats (red) interspersed with different variant repeats (purple, yellow, black, gray). At the other end of the amplicon there were 651bp corresponding to nucleotides 158242 to 158890 of the HHV6-A DR_R region (green) (GenBank: X83413.1). A primer called 10qF was designed based on the 10q sequence (blue arrow) and PCR was carried out with a primer specific for the DR8 region (DR8F green arrow) on genomic DNA from KUK and another cell line without CIHHV-6. This generated a product of ~1.1KB (C) at five annealing temperatures (58°C to 65°C) on KUK genomic DNA but not in a cell line without CIHHV-6 (WI38) confirming that the IPCR product corresponds to the internal junction.

The IPCR product was recovered and sequenced using nested primers (IPCR-1seq and IPCR-2seq see table 2.9 for primer sequences). The sequence revealed that the IPCR

product contained 28bp of unique sequence at one end that did not match with the consensus HHV-6A (strain U1102), with a restriction site of *Mbo*I at the very end of the molecule. This unique sequence was followed by 638 bp of TTAGGG repeats (and variants) and the final 651bp corresponded to nucleotide 158242 to 158890 of the HHV6-A DR_R region with some SNPs, but an overall identity of 96% (GenBank: X83413.1) (Figure 6.9B). In order to confirm the sequence represented the 10q-HHV6 junction in KUK, a primer was designed to the non-viral sequence (10qF) and PCR was carried out in conjunction with a primer specific for the DR8 region (DR8F). The PCR gave a product of ~1.1kb at 5 different annealing temperatures (58°C-65°C) and no products were obtained when DNA from a cell line without CIHHV-6 was used (Figure 6.9C). The PCR product was sequenced and it matched exactly with the sequence obtained from the IPCR product. These findings showed that the 28bp of unique sequence belonged to the 10q subtelomeric region in KUK. BLAST analysis of this sequence matched the 10q subtelomere, but also with 4q. This was not surprising since some subtelomeric regions share sequence similarity and it has been proposed some could have a common origin. This is known to be the case of 10q, 4q and 4p (van Geel M *et al.* 2002) (Riethman H 2008), To confirm that the 28bp belonged to the 10q subtelomeric region STELA was carried out using the 10qF primer with Teltail and Telorette2 on different cell lines including KUK and products were obtained in all of them (Figure 6.10). Taken together these data confirm that the IPCR product generated from KUK corresponded to the internal junction of the HHV-6A and 10q in this donor. Interestingly the breakpoint in the HHV-6A genome was similar to that described previously by Arbuckle JH *et al.*, with deletion of the terminal 431bp of the DR_R from the virus including the Pac2 element. However, unlike the junction described before (Arbuckle JH *et al.* 2010) the junction isolated from KUK contained several telomere variant repeats interspersed with the TTAGGG repeats. As the T2 region is thought to comprise only TTAGGG repeats the presence of telomere variant repeats likely represent the 10q telomere into which the HHV-6A integrated (complete sequence in appendix 7 section A7.5).

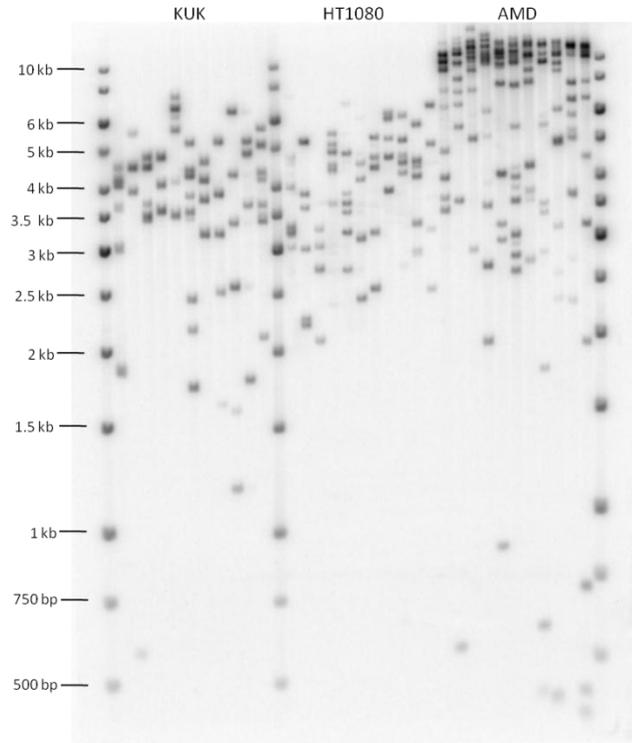


Figure 6.10 STELA Southern blot analysis in DNA from the KUK, HT1080 and AMD cell lines using the primer 10q. STELA PCR using the 10q primer (designed based on the sequence of the HHV-6-chromosome internal junction) and primers Teltail and Telorette2. STELA products were obtained in the three cell lines. This indicates that the 10q primer anneals within the 10q subtelomeric region. Genomic DNA was digested with *EcoRI* and 300 pg of digested DNA were used per reaction (per track). The products were detected following agarose gel size fractionation by Southern blot hybridisation to a ^{32}P radiolabelled telomere repeat probe. Sizes of the STELA products were calculated using the ImageQuant software (Amersham).

Attempts were made to isolate the internal junction from cell lines with B variant insertions (Table 6.8). The same IPCR approach followed by hybridisation to the HHV-6 and telomere probes was tried and a band of 1kb in all the cell lines and a ~2.2kb band in MP (Figure 6.11) were generated. The sequence of the 1kb band showed that it contained exclusively HHV-6 sequence, part was from the DR and part from the U1 region. This suggested that the circles were formed from the DR_L, which is closer to the telomere, instead of the DR_R (Figure 6.1). This explains why the fragment has the same size in all the cell lines. The 2.2kb fragment sequence from MP also comprised only HHV-6B sequence.

The difference in size of this fragment could be explained by the ligation and circularisation of two fragments after digestion instead of circularisation of a single fragment. The failure to amplify the fragment containing the internal junction could be due to the lack of an *MboI* restriction site on the chromosome close to the junction. After digestion the fragment containing the internal junction would be too long for IPCR amplification. In order to overcome this potential problem the sequence of DR8 from the DR_L was obtained to look for another restriction site. Assuming the identity of both DRs the same restriction sites should be present in the right repeat. Among the restriction sites present in the DR8 sequence the *AccI*, *AfeI* and *BsaI* endonucleases had a unique site (Figure 6.12). These enzymes were chosen to digest the genomic DNA of the cell lines harbouring CIHHV-6B prior to IPCR, unfortunately no products were observed after hybridisation, again possibly because of lack of a restriction site on the human chromosome close to the internal junction.

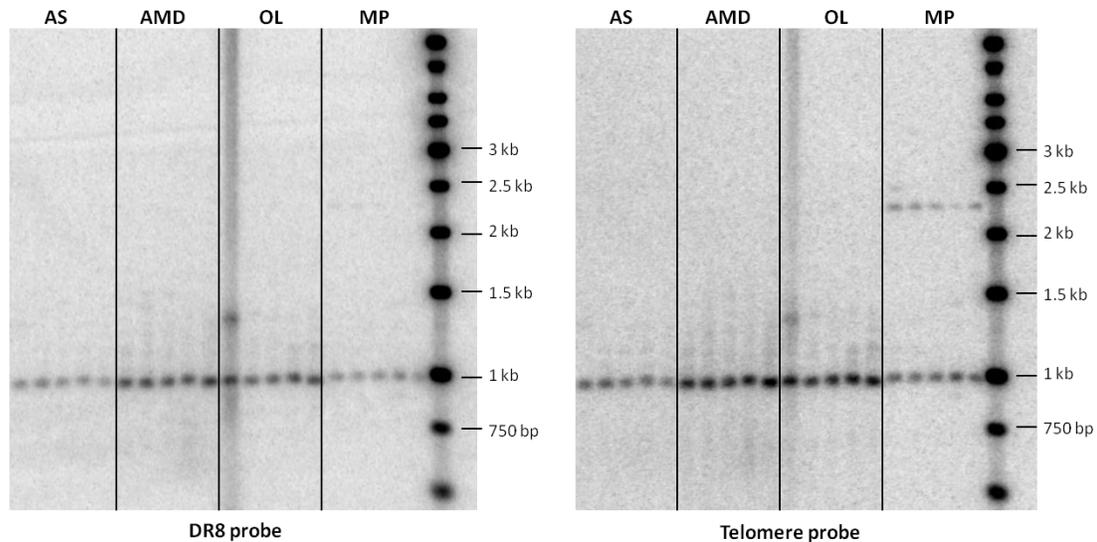


Figure 6.11 Southern blots of the IPCR products from the AS, AMD, OL and MP cell lines harbouring CIHHV-6B. Genomic DNA from the AS, AMD, OL and MP cell lines was digested with *MboI*. IPCR was carried out at five annealing temperatures (58°C-65°C). A 1kb product is present in all cell lines and in MP a second product of 2.2Kb was observed. Products were sequenced but none of them corresponded to the internal junction since only HHV-6B sequences were found.

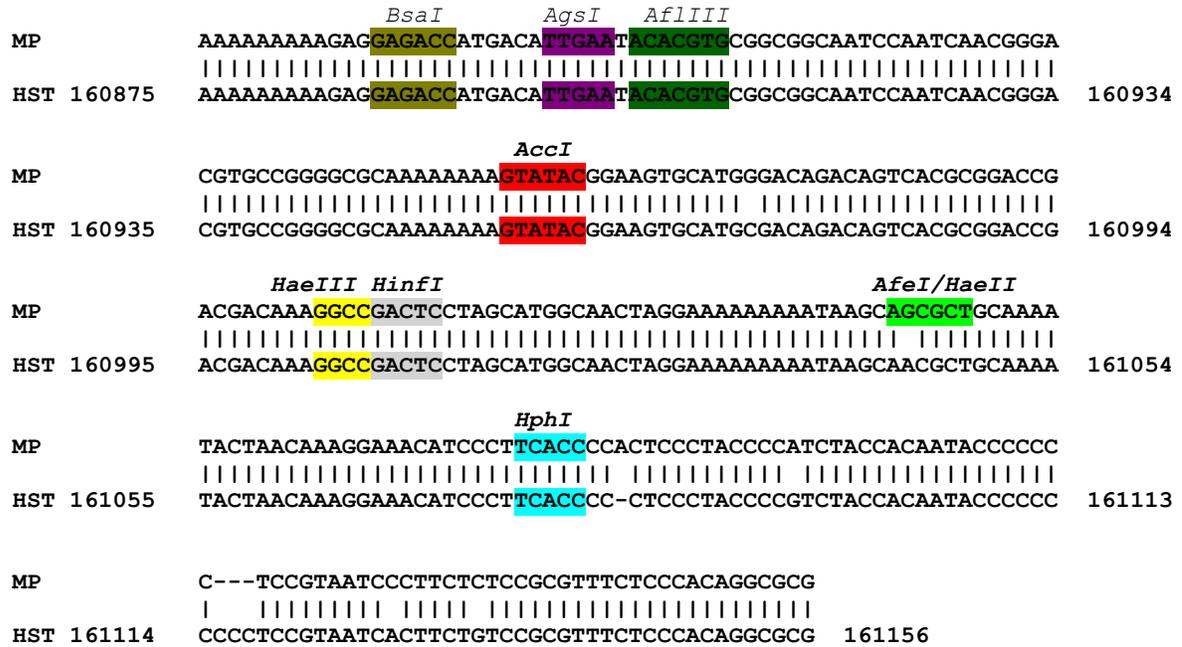


Figure 6.12 Alignment of the sequence obtained from DR8 region from the HHV-6B strain integrated in the lymphoblastoid cell line MP with the consensus sequence of the strain HST of HHV-6B (GenBank AB021506.1). The numbers indicates the nucleotide position according to the HST consensus sequence. The identity percentage between the sequences is 97%. All the restriction sites in the HST consensus sequence are present in the strain integrated in the MP cell line. In addition the MP cell line possesses an *AfeI/HaeII* restriction site. These restriction sites are upstream of the primers used for the IPCR.

Based on the knowledge that some subtelomeric regions share high sequence identity the primer 10qF was tested together with a DR8FU primer (see appendix 1 for sequence) which is different to the DR8F used previously because is able to anneal to the DR8 region of both HHV-6A and B. The PCR between the 10qF and DR8FU primers was done on genomic DNA from the AS, AMD, OL and MP cell lines plus another 13 lymphoblastoid cell lines obtained CIHHV-6 positive donors from the HGDP-CEPH screen, in which the integration site was unknown. KUK was used as positive control. A PCR product was only obtained in KUK (Figure 6.13). Currently other primers from subtelomeric regions of human chromosomes (Letsolo BT *et al.* 2009) are being used in attempts to obtain internal junctions.

Chromosome integrated human herpesvirus 6 (CIHHV-6) mechanism and consequences

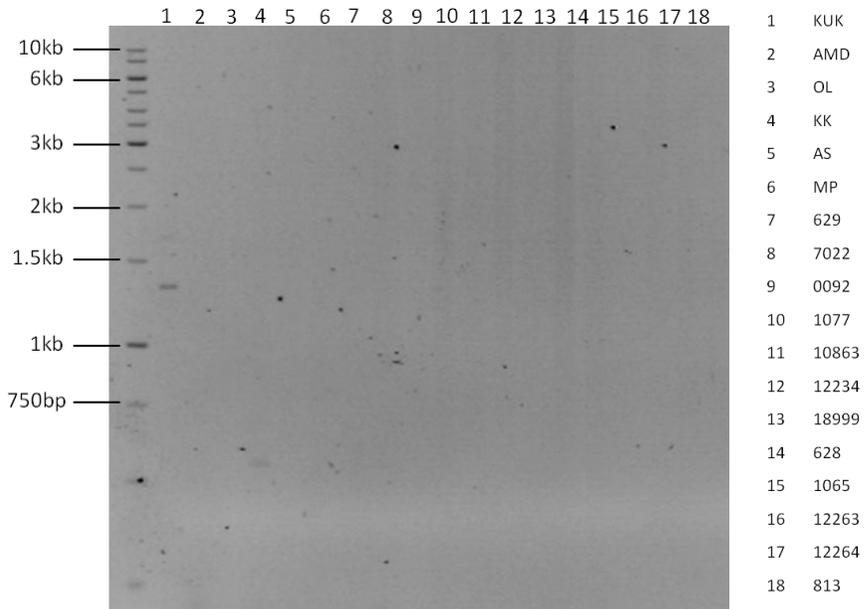


Figure 6.13 Agarose gel showing a PCR on genomic DNA from 18 different lymphoblastoid cell lines with CIHHV-6 using the primer 10qF designed from the DNA sequence of the proximal junction of KUK and the DR8F primer. The only positive was KUK, showing the expected product of ~1.3kb.

6.6.2 Sequence analysis of the distal breakpoint on chromosomes with CIHHV-6

The distal junction between the HHV-6 T1 element of the DR_L and the human derived telomere (Figure 6.1) was obtained from single STELA products. The sequence of the three siblings (1499, 1500 and 1501) was identical, confirming that the same virus is integrated in all of them. At its proximal end the sequence matched with nucleotides 686 to 162 of the HHV-6A consensus sequence (strain U1102, GeneBank: X83413.1), with an overall identity of 86%. Interestingly the identity from nucleotides 686 to 326 was 98% and the highest divergence was in the T1 element, which starts at nucleotide 338 in the consensus sequence. This divergence is mainly due to an expansion of the imperfect telomere repeats, which in this case span ~1kb. At the end of the molecule 348bp of pure TTAGGG were observed, with no evidence of the Pac1 element (Figure 6.14). This finding agrees with previous observations in Dr. Royle's laboratory (Gunjan Mukesh Wig *et*

al, unpublished data).

The sequence of sperm donor 1, which has CIHHV-6A was obtained, the results were similar. The overall identity of the sequence from nucleotide 686 to 162 was 86% with 98% identity before the start of the T1 element. There was an expansion of ~1.2kb of imperfect repeats. Surprisingly it was not possible to find a pure array of TTAGGG repeats at the end of the molecule (total length of the molecule 1.9kb), and again the Pac1 element was absent (Figure 6.14). Several attempts to sequence a longer molecule failed.

From the sperm donors with CIHHV-6B it was possible to obtain single STELA products to sequence only from donor 56. The inability to sequence STELA products from other donors was because most of the molecules obtained by STELA were longer than 4kb and it is difficult to re-amplify and sequence a molecule of this size. The length of the molecule sequenced from donor 56 was ~3kb. At one end the sequence matched with nucleotides 365 to 81 from the consensus sequence of the HST strain (GeneBank: AB021506.1) with an overall identity of 87%. This was followed by telomere like repeats (140pb) and then ~2kb of sequence that could not be sequenced due to the repeats. The final 500bp of the molecule still included telomere-like variant repeats without evidence of a pure array of TTAGGG repeats and again the Pac1 element was absent (Figure 6.14 for complete sequences see appendix 7 section A7.6). Similar results had been observed before in other samples with the CIHHV-6B (Dr. Royle *et al*, unpublished data). As mentioned before the T1 element is expanded in the CIHHV-6, particularly in the B type (up to 8kb). This may explain why a (TTAGGG_n) array was not found in the samples from the germline. It is possible that longer STELA molecules would reveal a pure (TTAGGG_n) array, however due to technical limitations it was not possible to sequence such long repetitive molecules.

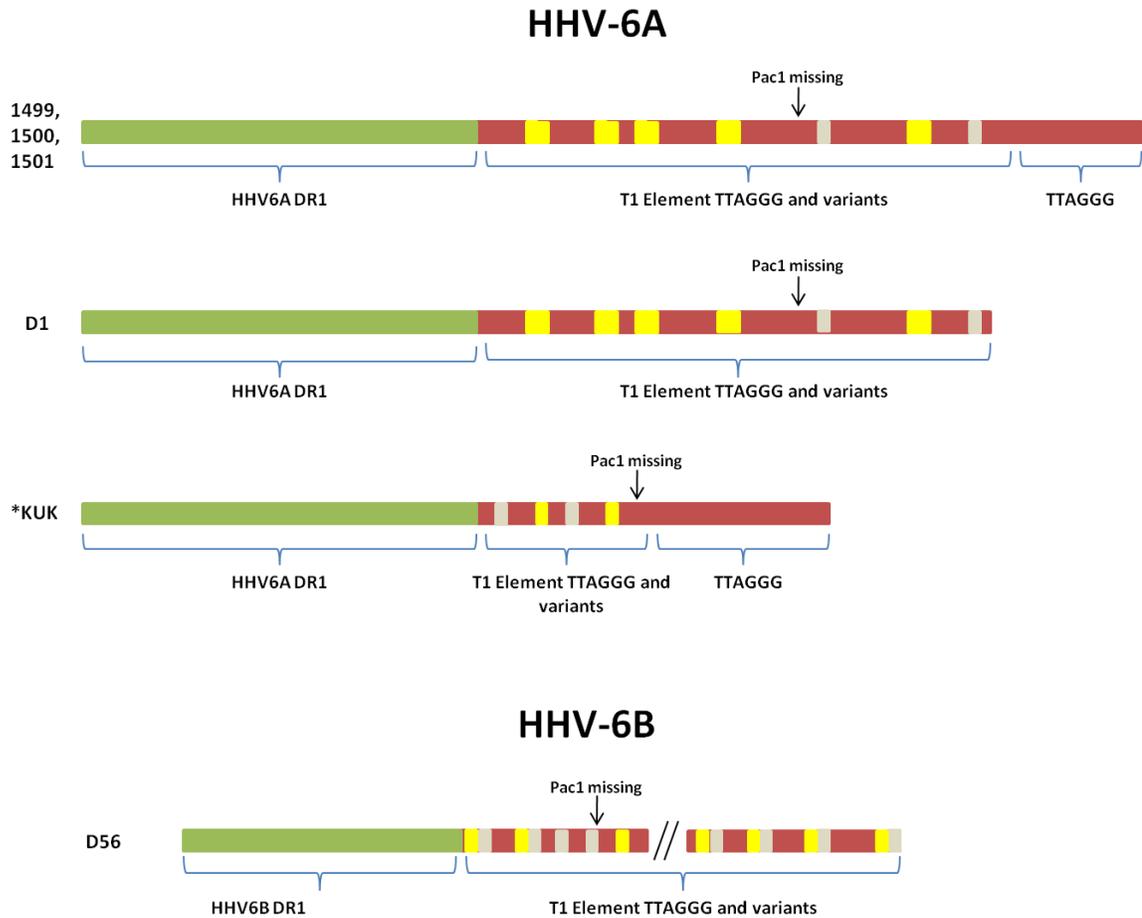


Figure 6.14 Schematic representations of the sequence organisation obtained from STELA products of the CIHHV-6 samples. In green is the sequence from the DR1 region, followed by the T1 element containing TTAGGG repeats (red) and variants (gray and yellow). A pure array of TTAGGG at the end of the telomere was found in samples from somatic DNA (1499, 1500, 1501 and KUK) but not in DNA from germline (D1 and D56). *KUK sequence organisation was obtained by Gunjan Mukesh Wig, *et al* unpublished data.

6.7 Discussion

After demonstration that HHV-6 integrates into human telomeres different groups around the world have studied the frequency of CIHHV-6 using a qPCR approach and the criteria proposed by Ward *et al*, which consider integration in individuals with $>6 \log_{10}$ copies/ml of whole blood (Ward KN *et al*. 2006). In a recent review a worldwide frequency of CIHHV-6 close to 1%, with a 1:3 ratio of HHV-6A:HHV-6B was reported (Morissette G *et al*. 2010).

It is important to bear in mind that not all published cases of CIHHV-6 have been confirmed with a different method like FISH, and it remains possible that some of the individuals had an active HHV-6 infection. In this study CIHHV-6 was determined by screening the samples for the presence of the U18 region and then confirming the positives by STELA to check for viral integration. The exceptions were the Spanish samples because the DNA from whole genome amplification is not suitable for STELA. The overall frequency obtained here is very similar to previous studies, however no positives were detected in African, American and Oceania populations. The absence of any HHV-6 positive samples in some populations may be due to the size of the sample, for example for Oceania just 28 samples were present in the panel. The 108 samples from America came from ethnic groups only from Mexico, Colombia and Brazil and although these are among the countries with the highest populations in the continent, the sample size from each population is small considering the large number of ethnic groups present in the whole continent. In addition research groups in United States of America (USA) have published studies reporting a frequency of ~1%. Some of these studies do not show ethnicity of the donors and some others divided the donors between white and non-white populations. The problem with this division of the populations is that due to the population diversity found in the USA, the non-white group can include Latin-American, African-American and Asian individuals. This makes it difficult to have clear picture of the frequency of CIHHV-6 in Latin-Americans, Afro-Americans or Asian populations living in the USA. Pellet *et al*, have reviewed these and other studies about frequencies reported by different authors (Pellett PE *et al*. 2011). Regarding the Africa sample, it is surprising that no HHV-6 positives were found. So far no surveys have been conducted to detect CIHHV-6 frequency in African countries, but some clinical trials have studied the incidence of HHV-6 in infant populations. Remarkably in the sub-Saharan population studied, HHV6-A accounted for 86% of the reported cases of HHV-6 infections (Bates M *et al*. 2009), whereas in other countries HHV-6B accounts for almost 100% of the cases, the reason for this divergence is not well known. Unfortunately there were no follow up studies or screening in parents to identify carriers of CIHHV-6. In contrast the Orkney and Sardinia

populations showed high frequencies, 4.7% and 7.1%, respectively but again this must be interpreted carefully because the sample size for each population was small. Ongoing work to screen more samples from Orkney will provide a more accurate frequency of CIHHV-6 carriers, nevertheless it is tempting to think that as both populations live on islands, historical isolation could have resulted in a founder effect and high frequency of CIHHV-6. Other factors influencing CIHHV-6 frequency variation among populations should be considered such as the virus interactions with the immune system. It is known that HHV6-A uses the CD46 cell surface protein as a receptor, and this is an important regulator of the complement system. So far no cell surface receptors have been identified for HHV-6B (Kawabata A *et al.* 2011). In addition, the responsible factors for the transportation of the virus into the nucleus are not known. It is possible that polymorphisms present in the CD46 gene, in the unknown cellular receptor used by HHV-6B or in the factors intervening in transporting the virus to the nucleus could modify susceptibility to infection or integration. This mechanism has been noted before, since some polymorphisms in the genes that play a role in apoptosis, (*FAS* and *FASL*), have been associated with a reduced susceptibility to Herpes simplex type 2 infection in African populations (Chatterjee K *et al.* 2010). In addition the HHV-6 U83 gene has a chemokine function over the group of CC-chemokine receptors that controls anti-viral leukocyte recruitment. Interestingly U83 is highly divergent between A and B subtypes this may contribute to their different pathogenicity. And it is possible that the group of receptors that respond to the U83 chemokine varies among populations, so this could modify the susceptibility to infection among individuals (Catusse J *et al.* 2009).

Michou V *et al.*, reported the presence of HHV-6 in semen indicating that it can reach the male germline, however the detected viruses were not integrated forms since they were not detected in the sperm samples after washing and centrifugation (Michou V *et al.* 2011). Here, the first report of CIHHV-6 frequency in DNA from the germline DNA is provided and the observed frequency in sperm samples from the British population was 5.4%. There are two possible explanations for the presence of CIHHV-6 in the male

germline: first is the inheritance of the CIHHV-6 from one of the parents, and second is the integration of HHV-6 after infection of the testis. The lack of somatic DNA or DNA from relatives of the sperm donors prevented further investigation. It is tempting to think that the difference in the frequency of CIHHV-6 between somatic and germline DNA could be related to transmission distortion. Transmission distortion is a phenomenon where an inherited character from a heterozygous progenitor does not follow Mendelian law of segregation (Meyer WK *et al.* 2012). Transmission distortion can result from two biological processes: the first is called “segregation distortion” and the second is viability selection. Segregation distortion includes meiotic drive (which means that a character is preferentially present in the functional gametes) and competition among gametes (which means that a character gives advantage to the gamete for fertilisation). Viability selection means that an inherited character could provide to the offspring a survival advantage or disadvantage (Meyer WK *et al.* 2012). To explore the possibility of transmission distortion the germline CIHHV-6 frequency should be studied further, first by enlarging the number of subjects and populations and second by trying to determine whether there is a disadvantage for the gametes carrying the CIHHV-6. An attempt to explore this possibility was done with the Spanish population samples. From the 671 samples analysed 171 belonged to infertile men and 500 to controls. I tried to investigate if the frequency of CIHHV-6 was higher in infertile Spanish men to establish whether there is a link between HHV-6 and male infertility that could explain the higher frequency in male germline. The frequency of CIHHV-6 was the same 0.58% and 0.59% for the group of infertile men and control Spanish men respectively. The other possibility to explain the higher frequency in the germline could be the viability selection. In this scenario an increased loss of embryos/foetuses carrying CIHHV-6 would be expected. Some studies have reported the presence of HHV-6 in cases of hydrops fetalis plus aneuploidies but it was not investigated if the virus was in an integrated form (Ashshi AM *et al.* 2000), (Al-Buhtori M *et al.* 2011). Further studies paying attention to the presence of CIHHV-6 need to be done to establish whether there is an association with pregnancy loss. In addition the study of large pedigrees of families with CIHHV-6 would provide more information about the rate of

CIHHV-6 transmission.

Previous telomere length analysis on DNA from the KUK cell line harbouring CIHHV-6A at 10q had shown that the HHV-6 associated telomere was the shortest, compared to 12q, 17p and XpYp telomeres. The findings in blood DNA from three siblings agreed with this observation, indicating an association between integration of the HHV-6 genome and short telomere length. The length analysis in germline DNA showed that the presence of the virus did not impede telomerase activity on the virus associated telomere as there was no evidence that this telomere was the shortest. When the number of molecules below 2.33SD was analysed in the four sperm donors no differences between the HHV-6 associated telomere and the 12q, 17p and XpYp telomeres were observed in donors 1, 44 and 56. The HHV-6 associated telomere in donor 32 showed significant differences compared with the other three telomeres. This observation should be interpreted carefully because the length analysis of the HHV-6 associated telomere in donor 32 showed two populations of telomeres. One of these populations is characterized by long telomeres (>10kb), it is possible that this population undergoes telomere trimming to regulate telomere length. This trimming mechanism has been described in germline cells (Pickett HA et al. 2011), and could generate a second population of telomeres characterized by a shorter length. The analysis of the number of molecules below 1.5kb showed that in some individuals the HHV-6 associated telomere was more prone to show truncation events, while in one of them there was no difference compared to the other telomeres analysed. Taking this data together it seems that CIHHV-6 does not interfere with telomerase activity, but the tendency towards reduced median telomere length in somatic cells suggests there may be a faster rate of telomere shortening. It will be necessary to explore other aspects like conformation of the subtelomeric chromatin. In the ICF syndrome, which is caused by mutation in the de novo DNA methyltransferase 3B gene (*DNMT3B*), patients possess short telomeres compared to healthy controls independent of the presence of telomerase activity. This seems to be related to hypomethylation of the subtelomeric regions leading a higher expression of telomeric-

repeat containing RNA (TERRA). A proposed model suggests that excess TERRA can interfere with replication fork progress by binding the template for the leading strand synthesis, the consequence of this would be an increase in telomere shortening (Yehezkel S *et al.* 2008). Therefore the methylation state and other epigenetic markers on the integrated virus genome, adjacent to the telomere, should be investigated to determine if it could influence the transcription of TERRA. It is possible that something similar could be happening in the germline since it is known that subtelomeric regions are hypomethylated in the germline and become methylated in the zygote (Brock GJ *et al.* 1999). In addition individual variation could also play a role. Based on twin and family studies it has been estimated that the heritability of mean telomere length is between 39% to 90%, but even when the heritability range is wide all authors agree in the importance of genetic factors in median telomere length and telomere loss rates (Monaghan P 2010). Furthermore linkage analyses have identified different loci associated with telomere length. The identified loci are 12p11, 14q23.2, 10q26.13 and 3p26.1. The authors have identified 10 candidate genes among all those loci that could be related to telomere biology. None of them correspond to telomerase components or shelterin complex members. However more work need to be done to establish a direct role of the identified genes on telomere length (Vasa-Nicotera M *et al.* 2005) (Andrew T *et al.* 2006).

The sequence analysis of the chromosome-HHV-6 proximal junction obtained from KUK revealed that the breakpoint in the viral genome found here is similar to those found in a family and clones of the HEK293 cell line by Arbuckle *et al.*, (Arbuckle JH *et al.* 2010). The T2 element comprises pure TTAGGG repeats, it is usually short but varies between strains (Achour A *et al.* 2009). No telomere variant repeats were described, in the chromosome-virus junction sequences described by Arbuckle *et al.* In the sequence they found different number of TTAGGG repeats after the beginning of the T2 element and then chromosomal sequence. Just one of the clones from the *in vitro* infected cells showed telomeric variant repeats within the TTAGGG repeats array. The sequence presented here from the 10q junction in the KUK cell line revealed 28bp of non-viral sequence which corresponded to

the 10q subtelomeric region, followed by 638bp of TTAGGG repeats interspersed with different telomeric variant repeats and finally sequence corresponding to the DR8 region from the DR_R of HHV6-A. It is known that the start of many human telomeres include sequence variant repeats before the pure (TTAGGG)_n. So it looks like the HHV6 found in KUK had integrated into the proximal region of a 10q telomere by HR with the T2 region of the DR_R of the virus. Also this sequence revealed loss of the HHV-6 Pac2 element as described before by Arbuckle *et al.*

The sequence from the HHV-6-telomere junction obtained from the STELA products, revealed the loss of the HHV-6 Pac1 element in the three siblings and the sperm donors. Furthermore, a pure (TTAGGG)_n array at the distal end of the telomere was found in the STELA products from the three siblings and KUK (Gunjan Mukesh Wig, *et al* unpublished data), but not in the STELA products from the sperm donors (D1 HHV6-A and D56 HHV6-B). This was unexpected since telomerase is active in the germline. One possibility is that these HHV-6 strains possess greatly expanded T1 regions, as has been shown recently in Dr. Royle's laboratory (Huang Y, *et al* unpublished data). The STELA products from the sperm donors that were sequenced were ~3.5 kb long. It is possible that sequences from longer molecules (~10kb) would show a pure array of (TTAGGG)_n at the distal end but technical limitations prevent the re-amplification and sequencing of such long molecules.

As stated before chromosomal integration is achieved by other herpesviruses. For example Epstein Bar virus (EBV) has been found integrated in some lymphoma patients and analysis in the Raji cell line (derived from Burkitt's lymphoma) have shown that EBV can integrate at different loci (Hurley EA *et al.* 1991), (Gao J *et al.* 2006). The other known example of herpesvirus integration is Marek's disease virus (MDV), which possesses regions of telomeric repeats and it has been shown that mutation of these repeats reduce integration efficiency (Kaufer BB *et al.* 2011). However telomeric repeats do not seem to be enough to mediate integration since the HHV-7 also has telomeric repeats but it has not yet been found in an integrated form. One important difference between HHV-6 and the other herpesviruses is the presence of the U94 gene which encodes a homologue of

the human adeno-associated virus type 2 (AAV-2) rep78/68 protein (Caselli E *et al.* 2007). It is possible that the U94 of integrated strains is different from that of non integrated, Arbuckle *et al.*, found 95% identity between the consensus sequence and the strains integrated in both families. Functional studies are needed to determine if such differences reflect a functional difference in the gene product (Arbuckle JH *et al.* 2010). In addition a model has been proposed in which integration is mediated by U94 after binding of the human telomere binding proteins TRF1/TRF2 to the T2 element (Arbuckle JH *et al.* 2011). To support this model it will be necessary to demonstrate that U94 is able to interact with members of the shelterin complex.

The idea that HHV-6 can achieve latency through integration into human telomeres and then be reactivated is under debate. The sequence analysis reported here demonstrated loss of the Pac1 and Pac2 elements, these elements are present in all the herpesviruses and are key factors for the cleavage and packaging of new viral particles (Tong L *et al.* 2010), How the loss of these sequences might affect the virus reactivation and generation of new functional viral particles is currently unknown. However Arbuckle *et al.*, have claimed the recovery of new viral particles after trichostatin A (which is an inhibitor of the class I and II of the histone deacetylases), treatment of cells carrying CIHHV-6 (Arbuckle JH *et al.* 2010). Even when the mechanism is not clear one possibility is that the virus folds itself to align the DRs and then through a HR-like mechanism between the T1 element an intact DR with Pac1 and Pac2 could be formed, and the virus could be released in a circular form. This would then allow replication and generation of concatamers harbouring the right signals for cleavage and packaging. Another plausible model is the formation of a t-loop structure by invasion of the telomeric single stranded 3' overhang into the T1 element. This structure would bring the DRs close and by an HR mechanism or t-loop excision the HHV-6 genome could be released (Figure 6.15). This is just speculative but it could explain how the virus can recover the Pac1 and Pac2 elements in the DRs, the homogenization of the DRs and the massive T1 expansion observed in the integrated strains of the virus compared with the non-integrated.

Chromosome integrated human herpesvirus 6 (CIHHV-6) mechanism and consequences

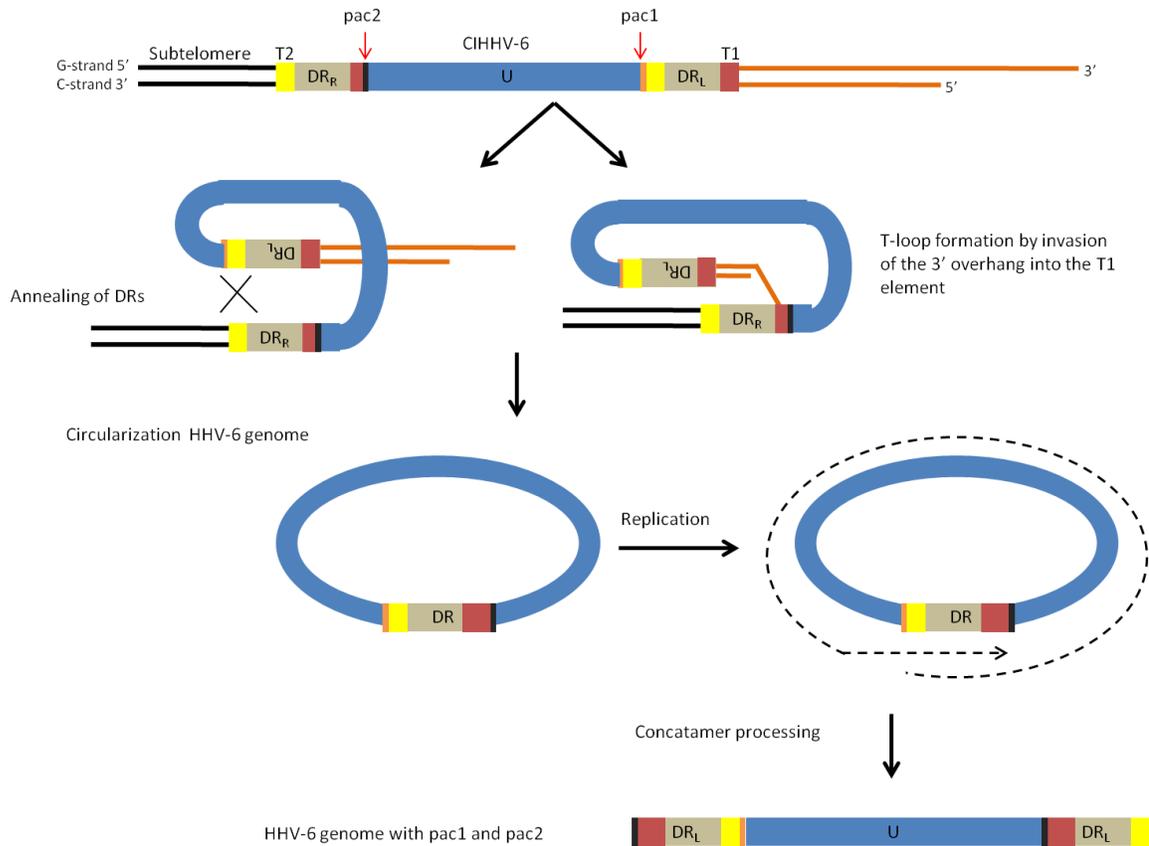


Figure 6.15 Possible mechanisms for CIHHV-6 reactivation. DRs are annealed by folding of the viral genome to promote HR between DR elements. Alternatively a t-loop structure is formed by invasion of the telomeric single stranded 3' overhang into the T1 region bringing the DRs close each other and then HR between the DRs or t-loop excision could happen. Both scenarios could result in circularisation and release of the HHV-6 genome from the chromosome with just one DR with an expanded T1 region and harbouring Pac1 and Pac2. This circle could be replicated by rolling circle replication and after concatamers processing new viral particles with both DRs and Pac elements could be generated.

6.8 Future work

HHV-6 is unique among the family of herpesviruses, its ability to integrate into human telomeres has opened a new field of research. Several groups are conducting studies to investigate the mechanism and consequences of CIHHV-6. The population screening conducted in this study have produced some interesting results, such as the high frequency in somatic DNA in donors from Orkney and Sardinia and the high frequency in sperm DNA from European/British donors. In addition this is the first report of telomere

length analysis in CIHHV-6 samples. The data provided here will help to improve the understanding of mechanisms of integration and the consequences on telomere function of CIHHV-6. Some of the further experiments that will contribute to the current knowledge of CIHHV-6 are:

- Screening of more samples from those population that showed high frequency of CIHHV-6. The CIHHV-6 frequencies found in the Orkney and Sardinia populations were higher than the worldwide frequency previously reported (1%). As stated before this could be due to sampling error so currently more samples from Orkney are being screened in Dr. Royle's laboratory. This will provide a more accurate frequency of CIHHV-6 in this population. Similarly it will be necessary to screen more samples from Sardinia to corroborate the frequency of CIHHV-6 found.
- Screening of more samples from germline from different populations. So far no data has been published regarding the frequency of CIHHV-6 in the germline. The CIHHV-6 frequency reported here (5.4%) is higher than the worldwide frequency reported for somatic DNA (1%). It is necessary to conduct a survey with more samples from the germline to verify the data and also to collect samples from different populations in order to make comparisons between them.
- Investigate transmission distortion. The higher CIHHV-6 frequency in DNA from germline compared to somatic DNA may suggest that transmission distortion could be happening. In order to investigate this possibility different approaches are needed. A) Family screening: According to the Mendelian segregation law, if one of the parents carries CIHHV-6, the expected CIHHV-6 frequency in the offspring should be 50%. Screening of large pedigrees of families carrying CIHHV-6 would confirm if the observed CIHHV-6 frequency is as expected. If not, this could be an indicator of transmission distortion. B) Screening of pregnancy loss: There are just a few studies reporting the presence of HHV-6 in pregnancy losses. However these studies did not specify whether HHV-6 was integrated or not. More studies are needed to establish whether CIHHV-6 can be a cause of pregnancy loss. This will

provide information about viability selection which can lead to transmission distortion. C) Sperm analysis. It would be interesting to investigate if 50% of the mature sperm population from a CIHHV-6 carrier harbours the CIHHV-6 as expected (meiotic drive). In addition it will be interesting to confirm that the sperm carrying CIHHV-6 have the same capacity to fertilise (competition among gametes). For this purpose fertilisation clinics evaluate morphology, motility and viability to select the optimum sperms.

- Epigenetic markers. The ICF syndrome is an example of how epigenetic modifications can be associated with telomere shortening. Investigation of the methylation levels of the integrated HHV-6 genome and the expression of TERRA would help to clarify whether the presence of the viral genome has an influence on the telomere length.
- DNA damage response. So far it is not known if the insertion of the HHV-6 genome into human telomeres triggers a DNA damage response (DDR). DDR can be monitored by detection of the phosphorylated form of the histone H2AX referred as γ H2AX. In order to investigate if CIHHV-6 triggers DDR it is necessary to look for localization of γ H2AX at the HHV-6 associated telomere and compare it with the other telomeres. Any increase in the DDR could be related to the trend of the HHV-6 associated telomere to be the shortest.
- HHV-6 latency-associated transcripts. It has been reported that some forms of the transcripts from the IE1 and IE2 regions are preferentially found in the latent state of the HHV-6 infection (Kondo K *et al.* 2002). The presence/absence of this so called HHV-6 latency associated transcripts (H6LTs) in CIHHV-6 cells will contribute to unravel whether the integration of HHV-6 into the human telomeres comprise the latent phase of its replication cycle.

CHAPTER 7 – FINAL DISCUSSION

Telomeres are specialized structures, composed of DNA and proteins, evolved to protect the integrity of human chromosomes thereby helping to insure genomic stability. One of the important factors for adequate telomere function is their length. Under normal circumstances, telomeres get shorter with each cell division, in the majority of somatic cells. Telomere attrition leads to cell senescence. If senescence is bypassed the cell can progress towards crisis when the genome becomes unstable and this usually ends in cell death. However, some cell lineages, such as the spermatogonium, need to actively extend their telomeres to maintain an adequate length. Telomere extension is accomplished by a ribonucleoprotein complex called telomerase. Telomerase expression is under strict control in most somatic cells. Nevertheless, some cells that bypass senescence successfully escape cell death by activating a telomere maintenance mechanism. Most cells re-activate telomerase but some activate a telomerase-independent mechanism called Alternative Lengthening of Telomeres (ALT). ALT is a homologous recombination-like mechanism and is found in 15% of human cancers.

Homologous recombination (HR) is a cellular mechanism used by cells to ensure accurate genome replication, DNA damage repair, meiotic recombination and chromosome segregation. HR is based on strand exchange between homologous DNA molecules (Wyman C *et al.* 2004). Current evidence suggests that homologous recombination (HR) can occur at telomeres. Early investigations on cells able to maintain their telomere length in the absence of detectable telomerase activity led to the discovery that sequences can be copied from one telomere to other (Dunham MA *et al.* 2000). Since then several evidence have emerged supporting this observation (Varley H *et al.* 2002). On the other hand, it has been shown that the human herpesvirus 6 (HHV-6) can integrate into human telomeres. HHV-6 includes telomeric repeats within the direct repeats of its genome. This observation has raised the possibility that the mechanism responsible for viral genome

integration is based on HR between the viral and the human telomeric repeats (Arbuckle JH *et al.* 2011). Here I have investigated these two phenomena involving HR at telomeres: the mechanism used by some cancer lineages to maintain their telomere length to avoid cellular death, and the integration of HHV-6 into the human telomeres.

7.1 The ALT pathway

The molecular mechanism underlying the ALT pathway remains elusive, but dissection of this pathway may allow the development of more specific and less harmful anti-cancer treatments, as well as the development of new biomarkers for diagnosis and prognosis. The genes involved in HR have been studied for possible roles in the ALT pathway. It has been demonstrated that the four genes studied here, *BLM*, *WRN*, *EXO1* and *CtIP*, play important roles in the DSB repair via HR throughout the genome. In addition functional interaction exist between these proteins (Nimonkar AV *et al.* 2008), (Nimonkar AV *et al.* 2011), (Eid W *et al.* 2010). The results obtained here demonstrated a lack of complex mutations in the ALT+ cell line W-V (*WRN*^{-/-}, that lack WRN). This suggests that WRN is important for promoting this kind of mutation which probably arise by copying or recombination with a non-homologous telomere or extra-chromosomal telomeric sequences. WRN is distinguished among the family of RecQ helicases because in addition to the helicase activity it possesses nuclease activity. Whether the helicase, the nuclease or both activities are needed to promote the formation of complex telomere mutations in ALT+ cells is yet to be determined. Furthermore, on reduction of BLM expression by 40% in the W-V cell line, an increase in telomere and MS32 mutation frequency was observed. These effects seem to be related to the ability of the WRN and BLM helicases to regulate strand invasion and resolution of HR intermediates (Wu L *et al.* 2006).

One of the critical steps for HR progression is the generation of long 3' tail for strand invasion and DNA synthesis (Holthausen JT *et al.* 2010). Based on experiments in yeast the exonuclease EXO1 is one of the main candidates for this process. Here attempts to

generate clones with downregulation of *EXO1* from two ALT cell lines were not successful. There could be different explanations for this observation. One possibility is that because *EXO1* is a multifunctional enzyme participating in DNA repair, its downregulation to low levels can cause cell death. It has been observed that in yeast strains with deletion of one of the components of telomerase the emergence of survivors was delayed when *Exo1* was mutated (Maringele L *et al.* 2004). Another possible explanation could be that downregulation of off-target transcripts by the shRNA used caused deleterious effects. However as two different shRNAs against *EXO1* were used and neither gave rise to clones with *EXO1* downregulation, off target effects seems an unlikely explanation. In addition, one of the shRNAs used (shEXOA) had been previously shown to downregulate *EXO1* expression by Bolderson E, *et al.* (Bolderson E *et al.* 2009). The other possibility is related to the vector and delivery method. More efficient transfection methods such as viral vectors should be tried to accomplish *EXO1* downregulation. In spite of the inability to obtain clones from ALT+ cell with *EXO1* downregulation, *EXO1* still looks like a good candidate to carry out resection of the 5' C-strand in ALT cells. This is supported by research showing that the resection of the 5' C-strand and consequent formation of 3' ssDNA tails at a DSB is affected by the deletion of *EXO1* (Nimonkar AV *et al.* 2011), (Mimitou EP *et al.* 2008), (Bonetti D *et al.* 2009).

The MRN complex associates with telomeres (Zhu XD *et al.* 2000) and depletion causes disruption of telomere maintenance in ALT+ cells but not in telomerase+ cells (Zhong ZH *et al.* 2007) with reduction in APBs formation in ALT+ cells (Jiang WQ *et al.* 2007). Experiments in human cell lines have demonstrated that the MRN complex interacts physically and functionally with CtIP. It has been shown that CtIP stimulates the nuclease activity of the MRN complex and CtIP depletion reduces the repair of DSB through HR (Sartori AA *et al.* 2007). In addition *Sae2*, the yeast homolog of CtIP, has a key role in the processing of the 5' C-strand at DSB. Experiments in yeast *in vivo* have suggested that there are at least two multi-protein machineries capable of processing the 5' C-strand at DSB generating long 3' ssDNA tails. One seems to be driven by *Sae2* and the other one by

Sgs1, the yeast homolog of BLM. Both machineries require the presence of an exonuclease, Dna2 and Exo1 seem to be the recruited exonucleases (Nimonkar AV *et al.* 2011), (Bonetti D *et al.* 2009)(Bonetti D *et al.* 2009, Nimonkar AV *et al.* 2011). It is tempting to think that something similar is happening at telomeres in ALT+ cells, so the effect of CtIP downregulation in ALT+ cells was investigated. The results obtained in this study showed for the first time that CtIP interacts with telomeric DNA in ALT cells much more frequently than in telomerase+ cells. This interaction could be happening at telomeres outside or inside APBs or with extra-telomeric DNA. To investigate further whether CtIP plays a role in telomere maintenance in ALT+ cells it would be interesting to identify if this association occurs at a particular phase of the cell cycle. Although there are contradictory opinions about the role of APBs in the ALT pathway, it would be helpful to investigate whether CtIP is present in APBs and if the association occurs at a specific phase of the cell cycle.

Two different shRNAs against *CtIP* (shCtIP) were cloned into a plasmid vector and two ALT+ cell lines were transfected with each of the shCtIP. After transfection, colony forming efficiency was analyzed, and the results showed that just one of the shRNA reduced the number of colonies in both ALT+ cell lines (shCtIPB). In addition, downregulation of *CtIP* was achieved in one clone from the ALT+ cell line SUSM1 after transfection with a plasmid carrying a shRNA against *CtIP*. This clone (shCtIPA1) showed approximately 50% reduction of CtIP expression at PD19 and PD21. The SUSM1 shCtIPA1 clone showed reduced growth rate when compared with controls. Despite the reduced growth rates, no effects were observed in the markers of ALT that were analyzed. Interpretation of the data obtained here could be that CtIP does not play a role in the ALT pathway or that 50% downregulation of *CtIP* expression is not enough to affect the markers of ALT studied. Alternatively perhaps no effect was detected because of the existence of another protein or proteins able to compensate for the reduced levels of CtIP. This would be something similar to what has been observed in yeast, where Sae2 and Sgs1 act through independent routes to direct the DSB repair as mentioned above (Bonetti D *et al.* 2009), (Nimonkar AV

et al. 2011).

It is attractive to think that there could be two routes for the processing of the 5' C-strand and progression of HR between telomeric sequences in ALT+ cells. On the one hand it is possible that following recognition of a critically short telomere, the MRN complex in conjunction, with CtIP arrive at the telomere and initiates the resection of the 5'C-strand and then an exonuclease such as EXO1 or DNA2 is recruited and long 3'ssDNA tails generated. On the other hand it is possible than BLM in association with EXO1 or DNA2 direct the efficient processing of the 5' C-strand to generate long 3' ssDNA tails. Future work is needed to determine if these models are plausible and understand the regulation of these mechanisms.

7.2 Chromosomally integrated HHV-6

HHV-6 is the causal agent of exanthema subitum in children and since it was found in patients with lymphoproliferative disorder it has been proposed as a risk factor for cancer development. Now it is well document that HHV-6 integrates into human telomeres. Chromosomally integrated HHV-6 (CIHHV-6) represents an unusual behaviour for members of the family of herpesviruses and the mechanism and consequences of integration are not known yet. The worldwide frequency of CIHHV-6 is 1% (Morissette G *et al.* 2010), however the available data have been obtained from a limited number of populations. This study screened population from around the world and compared their CIHHV-6 frequencies. Some population showed interesting results, like the high frequency in Orkney and Sardinia or the absence of positive samples in the populations from Africa and America. Whether this is related to genetic factors specific for each population or to the virus strains present in a specific part of the world is yet to be determined. Current work in Dr. Royle's laboratory is aiming to generate a phylogenetic tree to understand whether the viruses found in an integrated form are more closely related to one another than to non-integrated forms. In addition more samples from the Orkney population are

being studied to corroborate the integration frequency observed.

HHV-6 has been found in semen, but not in an integrated form (Michou V *et al.* 2011). This is the first report of the frequency of CIHHV-6 in the male germline. The frequency in DNA from sperm from European/British cohort was 5.4% and when compared with the frequency found in somatic DNA it is statistically significant different. More studies are needed to investigate if this difference is simply due to a sampling error or if there is a selective disadvantage. For example it is possible that gametes carrying a CIHHV-6 have a disadvantage for fertilization. Or the zygote produced by fertilization with a gamete harboring a CIHHV-6 could suffer disadvantage. In addition, telomere length analysis by STELA in DNA from the CIHHV-6 positive sperm donors showed that the presence of the viral genome did not compromise telomere extension by telomerase in the male germline. Furthermore, the analysis of critically short telomeres in the male germline did not show strong evidence that solely the presence of CIHHV-6 can increase the proportion of short telomeres.

This study provides the first analysis comparing the length of the HHV-6 associated telomere with other three telomeres in DNA from blood from 3 siblings and from lymphoblastoid cell lines, all of them with confirmed CIHHV-6. The results suggest a trend towards the HHV-6 associated telomere often being the shortest among the studied telomeres. This observation needs more investigations and work in Dr. Royle's laboratory in analyzing more lymphoblastoid cell lines with CIHHV-6 is being conducted to verify this observation. In addition the possible association with CIHHV-6 and DNA damage response will be explored.

Previous analysis of the chromosome-HHV-6 proximal junction had shown absence of the Pac2 element in the DR_R of the virus (Arbuckle JH *et al.* 2010). The sequence obtained from the chromosome-HHV6 proximal junction, from the KUK cell line with CIHHV-6 at 10q, reported here also showed loss of the Pac2 element. In addition the sequence obtained from the chromosome-HHV6 distal junction from the 3 siblings and two sperm

donors showed the loss of the Pac1 element in the DR_L of the virus. The loss of the Pac1 and Pac2 elements could happen during the integration of HHV-6 through homologous recombination between the human telomere and the perfect telomere repeats present at the T2 element in the DR_R in the viral genome.

The debate about whether CIHHV-6 is a form of viral latency is still going on and two possible mechanisms for the release of the virus as a circular DNA molecule have been proposed here (Figure 6.12 chapter 6). Current work in Dr. Royle's laboratory is trying to identify the presence of circular forms of the HHV-6 genome in DNA from the lymphoblastoid cell lines with CIHHV-6. Equally important is to investigate if reactivation from an integrated form occurs *in vivo*, and identify factors, such as drugs, that could trigger the release of the HHV-6 from the integrated form.

The field of CIHHV-6 research is relatively new and requires a multidisciplinary approach to identify all the risks that accompanying it. This will allow the creation of health schemes to survey and protect the population.

APPENDICES

APPENDIX 1 Solutions and buffers

11X PCR buffer

Reagent	Stock conc	Conc in 11.1X	Conc in final reaction
Tris HCl pH 8.8	2M	500mM	45mM
Ammonium Sulphate	1M	122mM	11mM
MgCl₂	1M	50mM	4.5mM
2-mercaptoethanol	100%	74mM	6.7mM
EDTA pH 8.0	10mM	49uM	4.4uM
dATP	100mM	11.1mM	1mM
dCTP	100mM	11.1mM	1mM
dGTP	100mM	11.1mM	1mM
dTTP	100mM	11.1mM	1mM
BSA Ambion	50mg/ml	1.25mg/ml	113ug/ml

Oligonucleotide labelling buffer (OLB)

10µl Solution A

25µl Solution B

15µl Solution C

Solution A

1250µl 2M Tris HCL pH 8.0

50µl 5M Magnesium Chloride

700µl Elga water (double deionised/distilled)

36µl 2-Mercaptoethanol

10µl dATP - from 100mM Stock

10µl dGTP - from 100mM Stock

10µl dTTP - from 100mM Stock

Make in a 2ml Eppendorf tube, mix thoroughly, aliquot in 500µl lots. Store in -20°C freezer.

Solution B

2M Hepes pH 6.6

Made from Hepes (free acid) and aliquoted in 1ml amounts. Store in -20°C freezer.

Thaw 1 aliquot for each batch of OLB.

Solution C

Pd(N)₆ Random Hexamer Sodium Salt (from Fisher/ Fermentas product no. SO142) Comes

as a 120 μ l stock containing 24ug of Hexamer.

Formamide loading dye

80% formamide

10mM EDTA pH 8.0

1mg/ml xylene cyanol FF

1mg/ml bromophenol blue

APPENDIX 2 pSUPERIOR plasmid

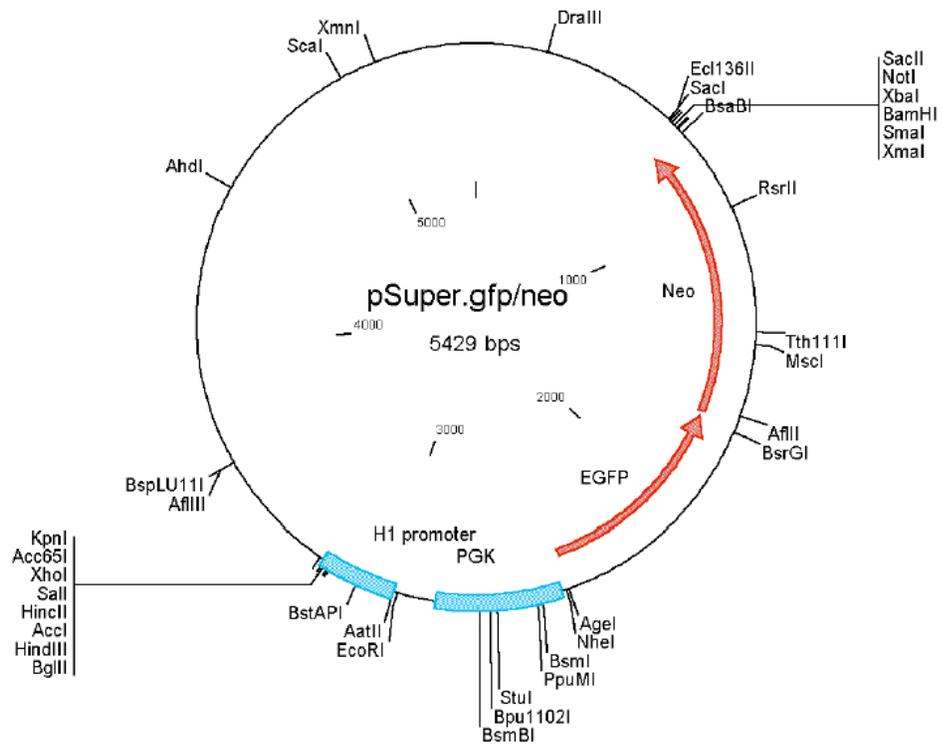
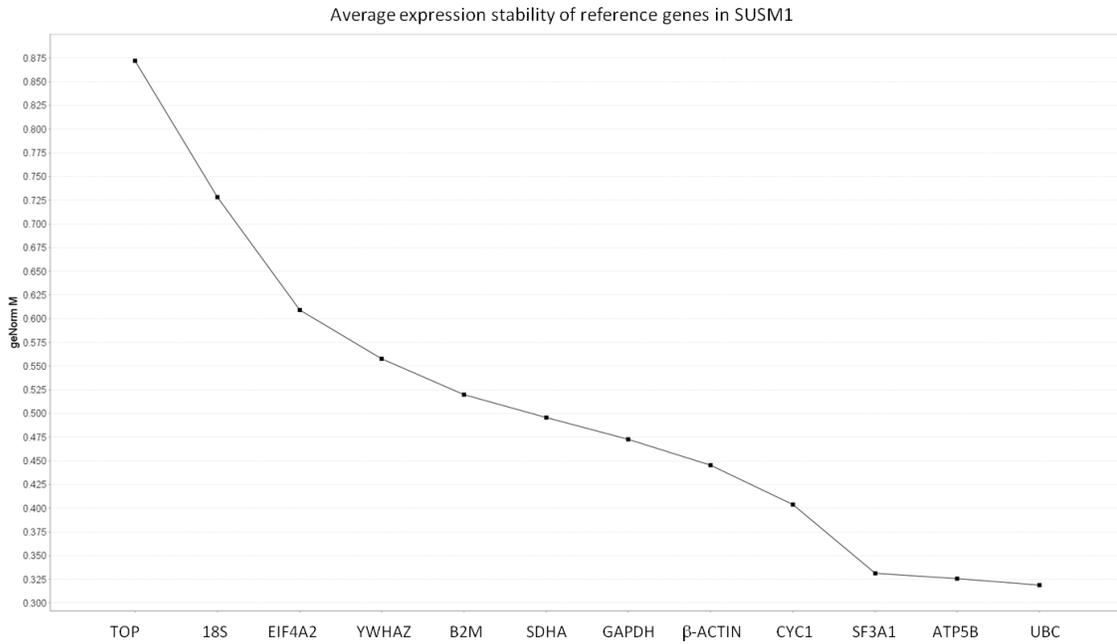


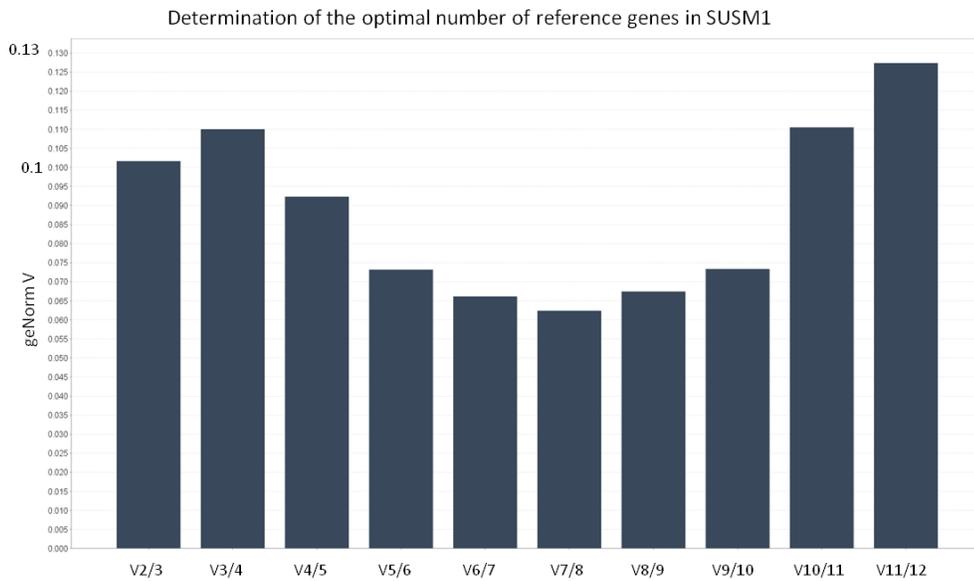
Figure A.1 Plasmid pSUPERIOR.neo+GFP. Length 5430 bp, key sites BglII 3182, HindIII 3188, EcoRI 2960, Sall 3203, XhoI 3209. H1 promoter 2965-3167, neo ORF 1684-715. PGK promoter 2840-2442, egfp orf 2424-1691, ampicillin resistance cassette ORF 5302-4445.

APPENDIX 3 Selection of reference genes for real time PCR

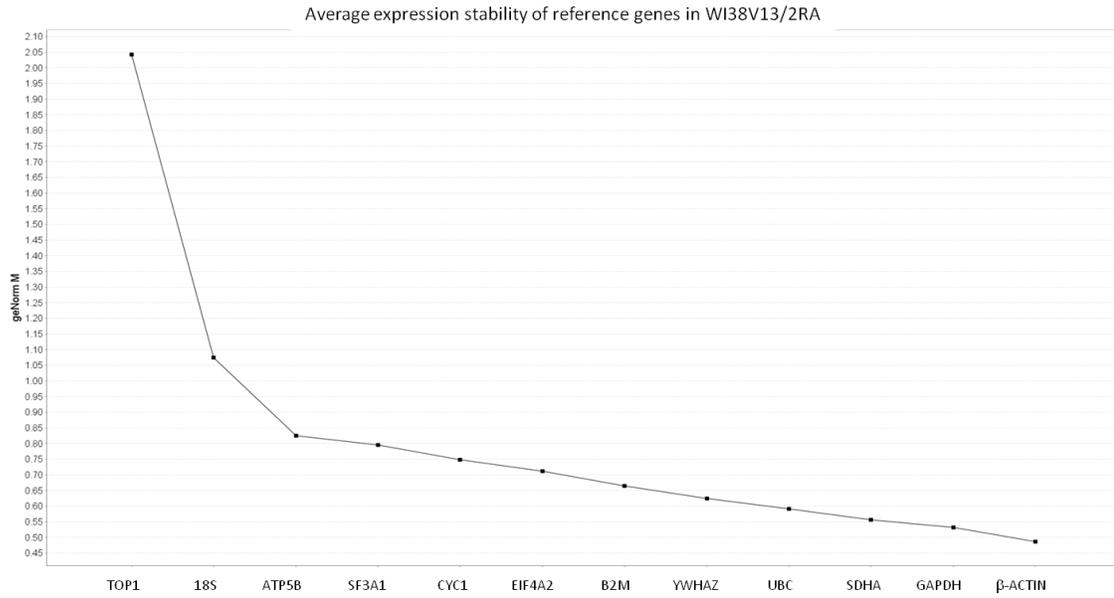
A



B



C



D

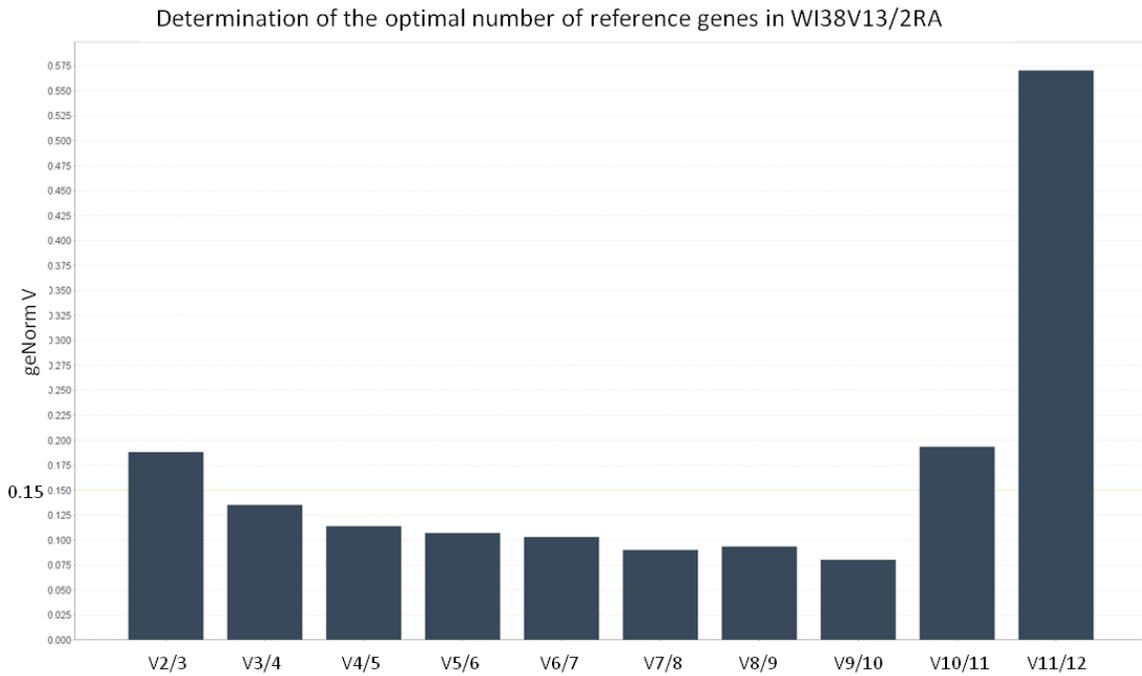
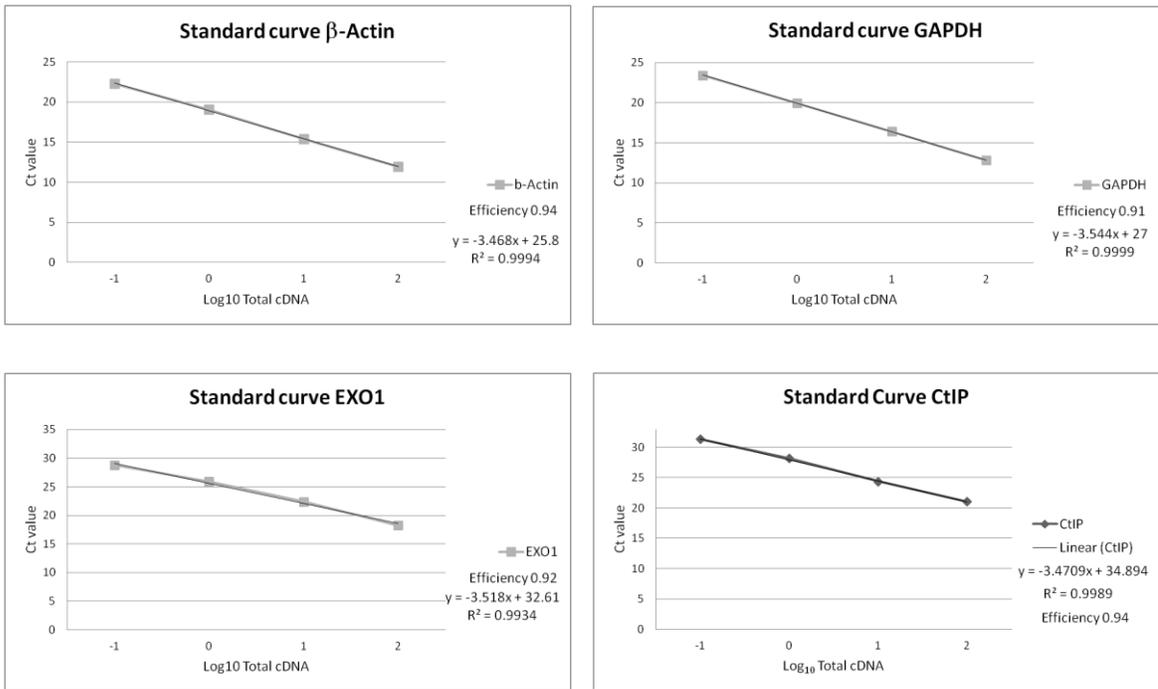


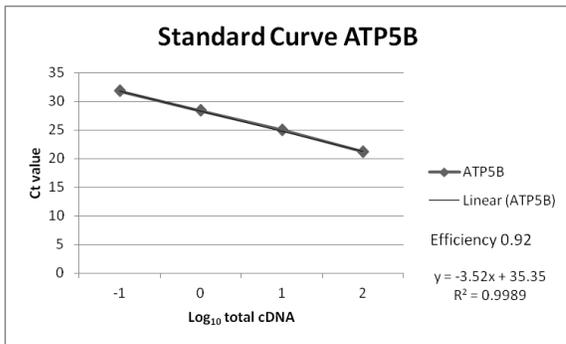
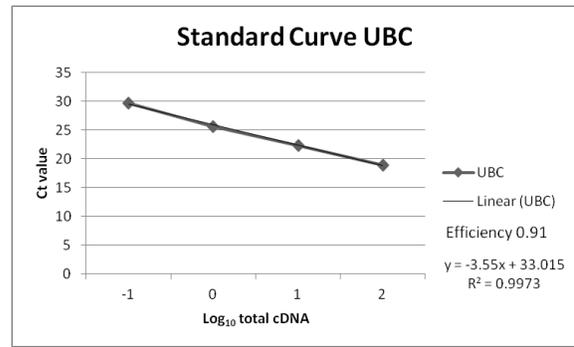
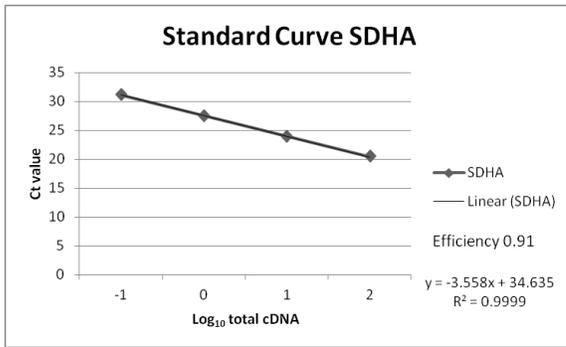
Figure A.2 Determination of the reference genes with the most stable expression and the number of reference genes

needed for adequate normalisation of the qPCR data in the ALT+ SUSM1 (A and B) and WI38V13/2RA (C and D) cell lines using the geNorm applet included in the qbase^{plus} software (Biogazelle). The 12 reference genes contained in the geNorm reference gene selection kit from Primer design were tested on eight different cDNAs from eight different cell pellets of each cell line. (A and C) The geNorm applet compares the expression of each gene against the others. It makes all the possible gene combinations within each sample and then across all the samples. The value obtained is called M value and as it correlates inversely with the gene expression stability. The genes on the right of the graph are the ones with the lowest M value and therefore with the most stable expression for the SUSM1 cell line. (B and D) The geNorm applet determines how many genes are needed to obtain an adequate normalisation factor for the qPCR data of each cell line. The first column from the left (V2/3) represents the use of two of the three genes with the most stable expression, the second column (V3/4) represents the use of three of the four genes with the most stable expression. The applet adds one gene each at the time. The normalisation factor is evaluated with the V value, the number of genes giving a V value below 0.15 is considered adequate for adequate normalization of the qPCR data. In the case of SUSM1 the two genes with the most stable expression (UBC and ATP5B) gave a V value <0.15 (B). These two genes are needed for normalisation of the qPCR data. In the case of WI38V13/2RA the three genes with the most stable expression (b-ACTIN, GAPDH and SDHA) gave a value <0.15 (D). These three genes are needed for normalisation of the qPCR data. Graphs obtained from the qbase^{plus} software. All reactions were done in triplicate.

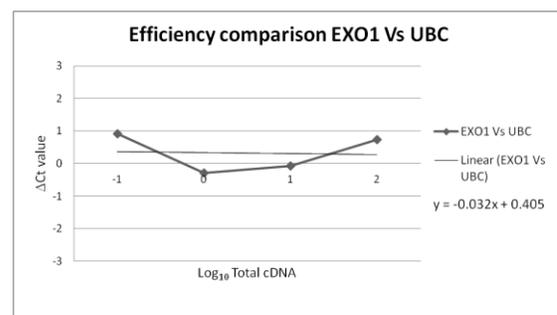
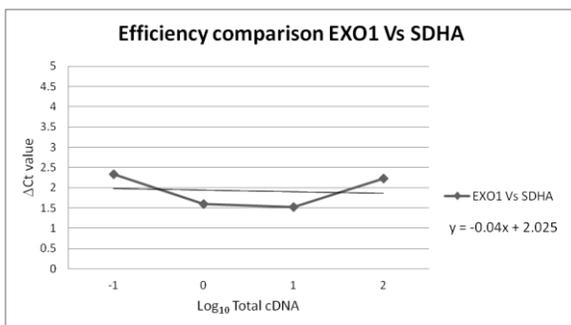
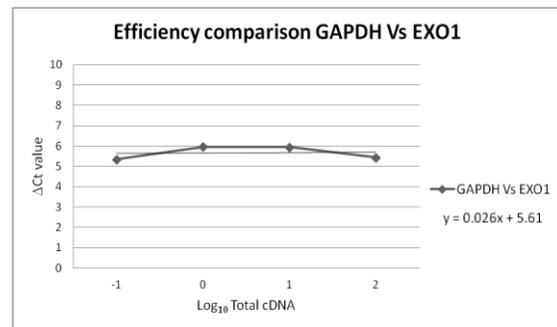
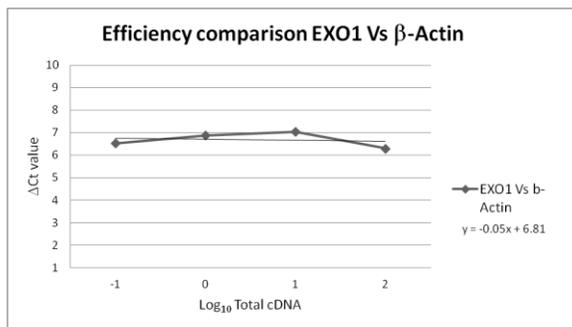
A3.1 Real time PCR efficiencies

A





B



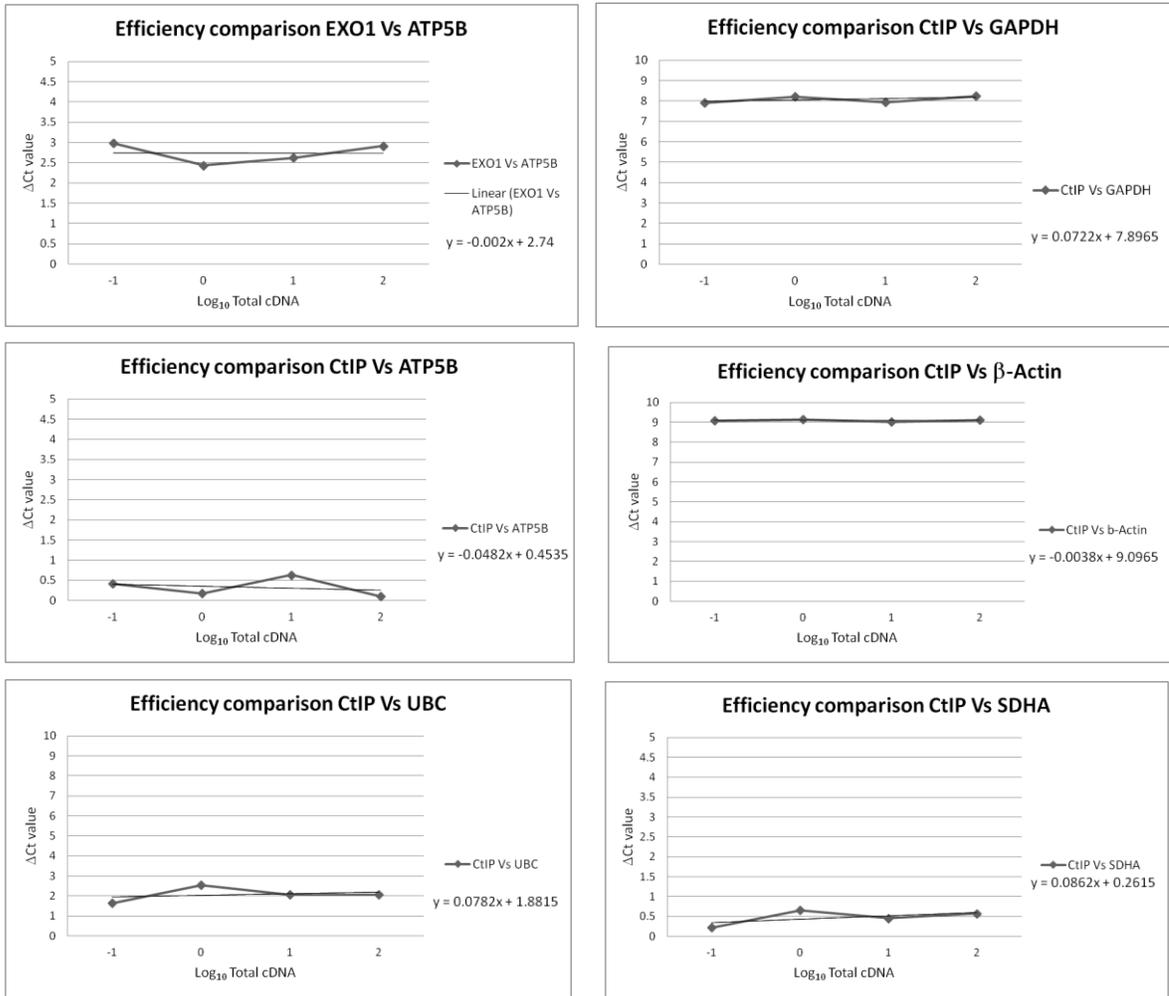


Figure A.3 Standard curves to calculate the amplification efficiency of the β-Actin, GAPDH, EXO1, CtIP, UBC, SDHA and ATP5B amplicons and comparison of the amplification efficiency between amplicons. (A) Standard curves were generated using serial dilutions of cDNA as described in section 2.11.2. On the Y axis are the Ct values and on the X axis the log₁₀ of the amount of cDNA used at each point of the dilution series. On the right of each graph is the efficiency of the qPCR, calculated using the value of the slope obtained from the equation of the linear regression displayed below the efficiency value (A). (B) Comparison between amplicon efficiency (reference genes Vs target genes) is necessary to validate the use of the selected reference genes. This graph plots the ΔCt value (ΔCt = Ct target gene – Ct reference gene) on the Y axis against the log₁₀ of the amount of cDNA used at each point of the dilution series on the X axis. Linear regression equations are displayed on the right of each graph. A value of the slope <0.1 indicates that the efficiencies are very similar and so the reference gene can be used. All reactions were done in triplicate.

Appendix 4 *EXO1* expression analysis in SUSM1 clones

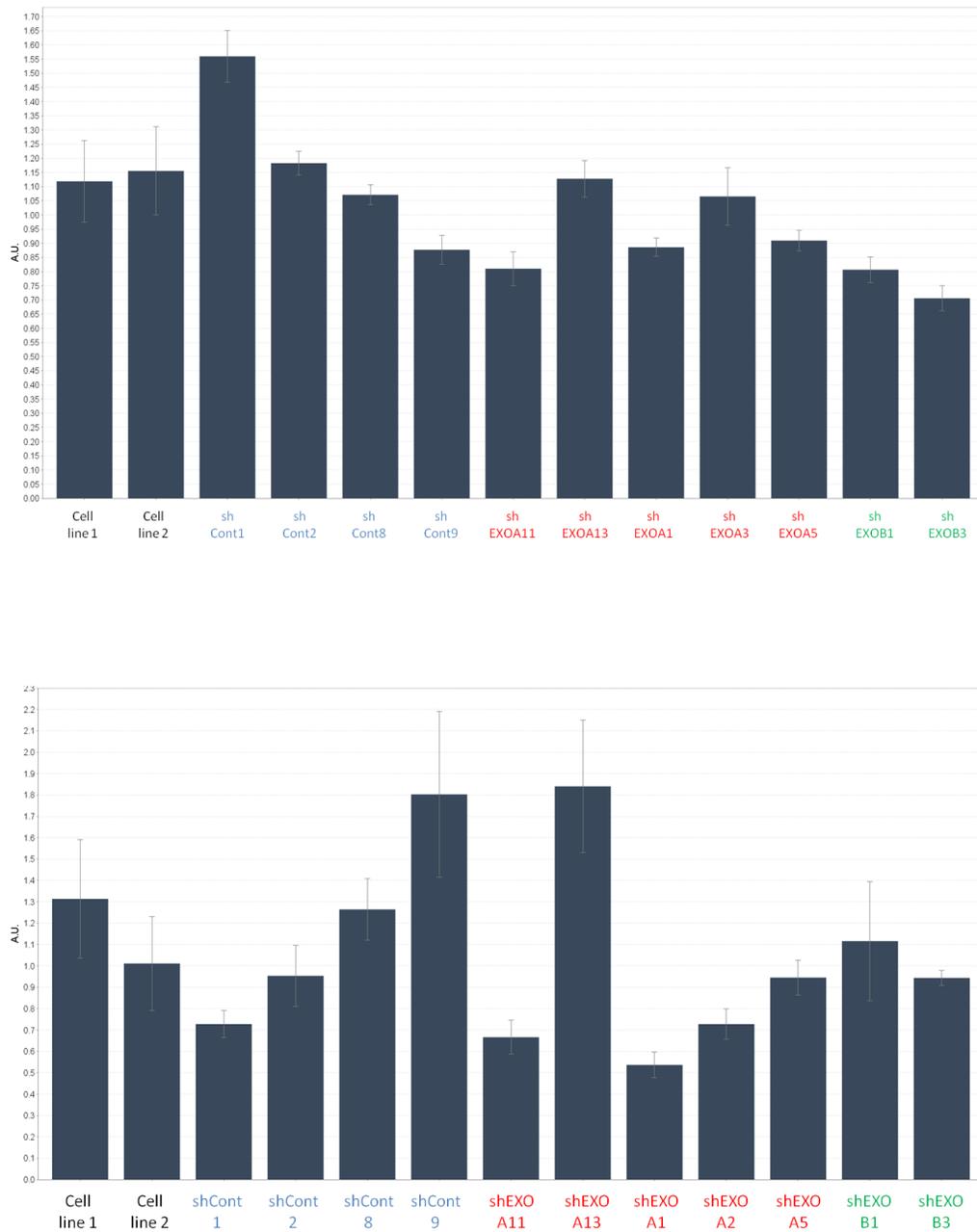


Figure A.4 Replicas of the *EXO1* expression analysis by qPCR in the SUSM1 clones expanded after transfection of the SUSM1 cell line with a vector containing a shRNA against *EXO1* or an empty vector as control. Comparison of the *EXO1* expression analysis between the progenitor SUSM1 cell line, control clones and shEXO A and shEXO B clones. All qPCR reactions were set up in triplicate. There is no evidence of downregulation of *EXO1* expression in any of the SUSM1 clones. Data was normalised using ATP5 and UBC as reference genes using the qbase^{plus} software. A.U. arbitrary units.

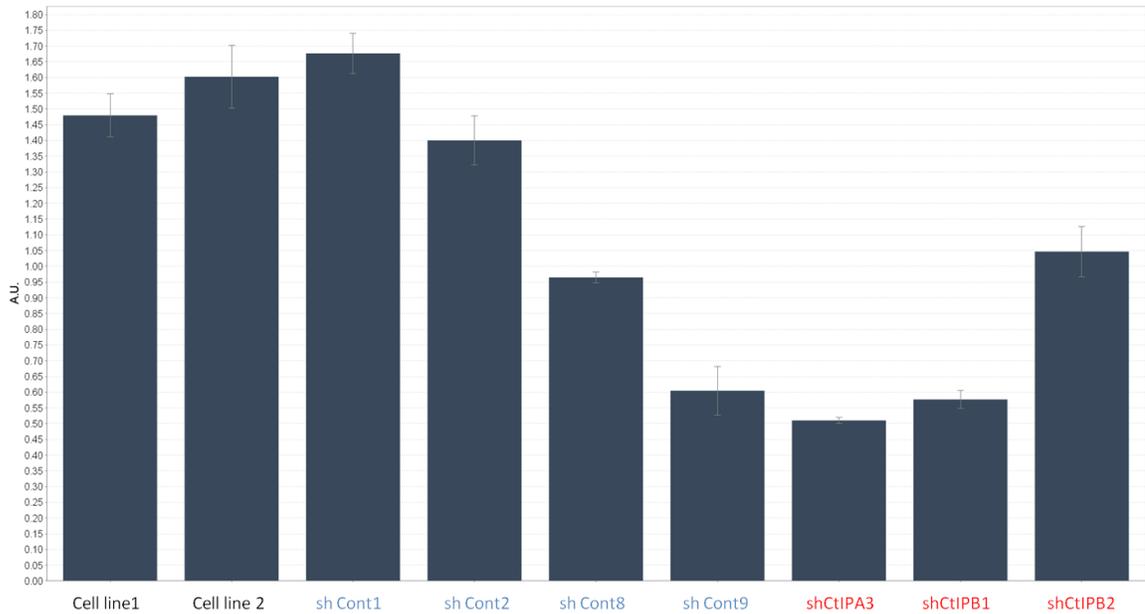
APPENDIX 5 *CtIP* expression analysis in SUSM1 clones

Figure A.5 Replica of the *CtIP* expression analysis by qPCR in the SUSM1 clones expanded after transfection of the SUSM1 cell line with a vector containing a shRNA against *CtIP* or an empty vector as control. Comparison of the *CtIP* expression analysis between the progenitor SUSM1 cell line, control clones and shCtIPA and shCtIPB clones. All qPCR reactions were set up in triplicate. There is no evidence of downregulation of *CtIP* expression in any of the SUSM1 clones. Data was normalised using ATP5 and UBC as reference genes using the qbase^{plus} software. A.U. arbitrary units.

Appendix 6 Analysis of the WI38V13/2RA clone shCtIPA1 and controls.

A6.1 Southern blot of Single telomere length analysis (STELA) of the WI38V13/2RA clone shCtIPA1 and controls.

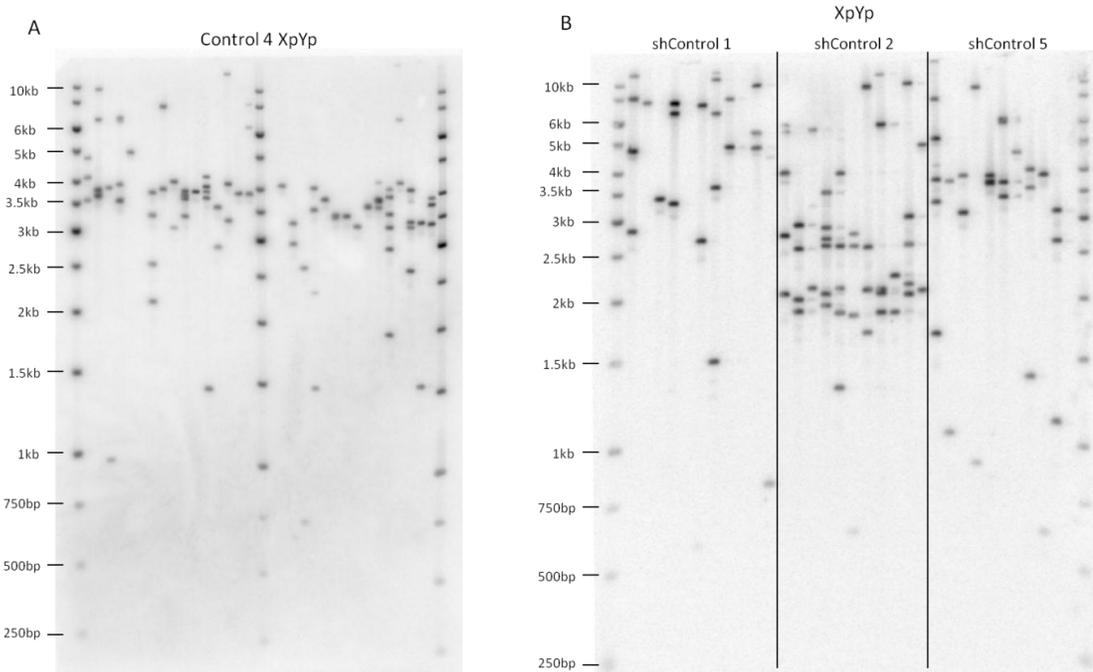


Figure A.6 Southern blot of XpYp STELA from WI38V13/2RA control clones after hybridisation with a ^{32}P labelled telomere probe. (A) Shows control 4 and (B) controls 1, 2 and 5. Unlike the WI38V13/R2A shCtIPA1 clone (Figure 5.6) all the 4 controls show the classical telomere length heterogeneity expected for an ALT+ cell line. PCR reactions contain 300 pg of *EcoRI* digested genomic DNA.

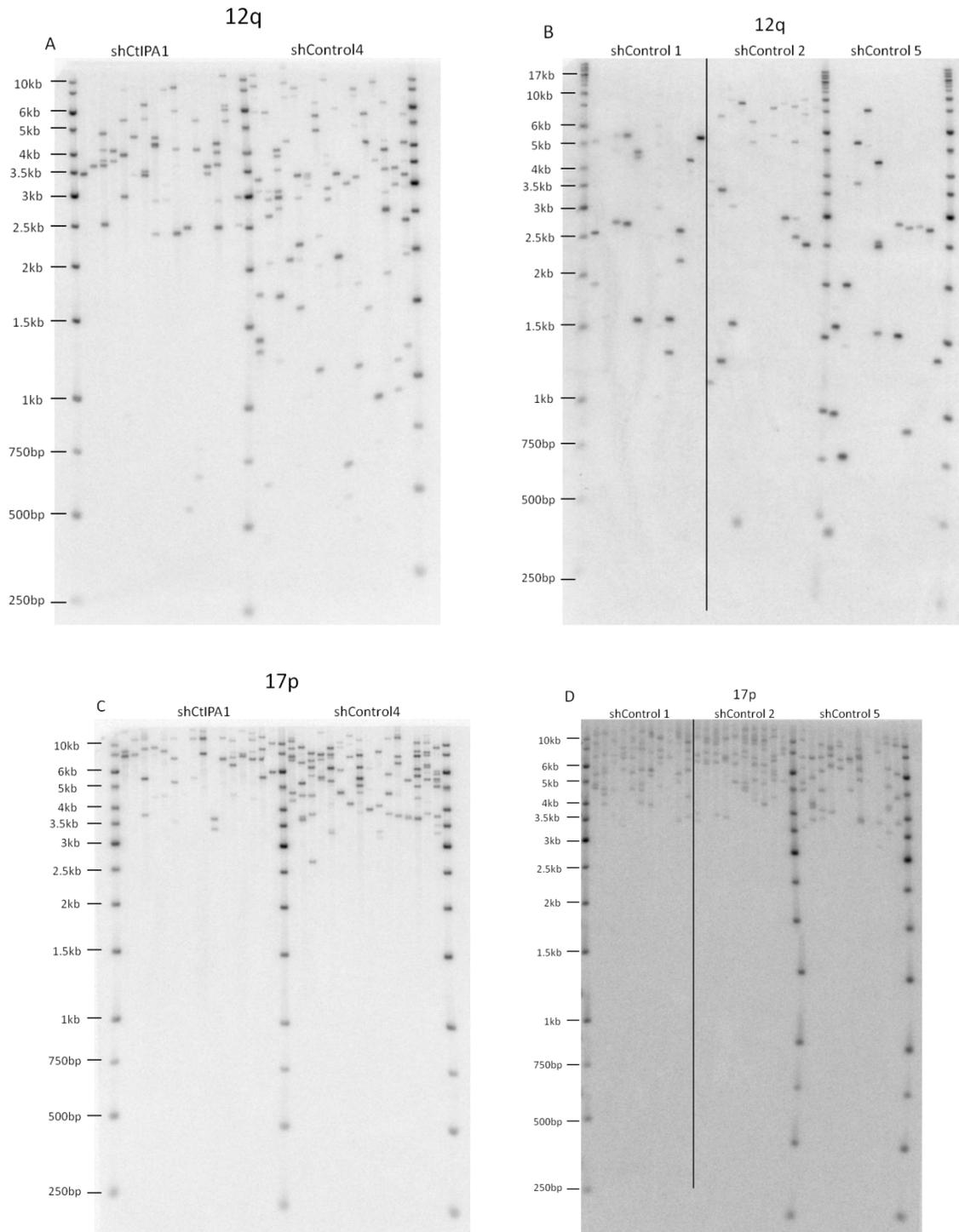


Figure A.7 Southern blot of STELA from WI38V13/2RA shCtIPA1 and control clones after hybridisation with a ^{32}P labelled telomere probe. A and B show length distribution of telomere 12q. C and D show length distribution of telomere 17p. All the studied WI38V13/2RA clones show the classical telomere length heterogeneity expected for an ALT+ cell line. PCR reactions contain 300 pg of *EcoRI* digested genomic DNA.

A6.2 C-circle quantification in the WI38V13/2RA cell line, clone shCtIPA1 and controls

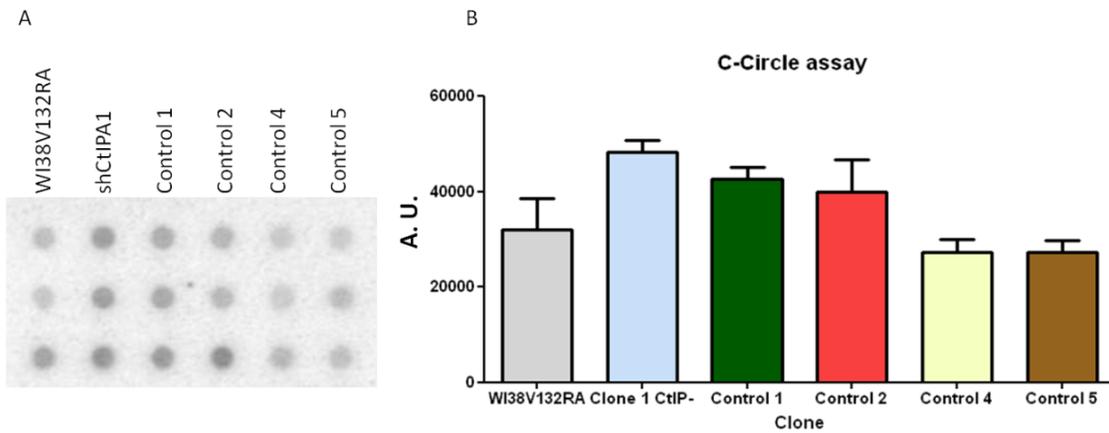


Figure A.8 Replica of the quantification of telomeric DNA generated in the C-circle assay from the WI38V13/2RA cell line, shCtIPA1 and control clones. (A) Dot blot of the C-circle amplification reactions after hybridisation to a ³²P labelled telomere probe. Reactions were done in triplicate. Each reaction contained 10 ng of *EcoRI* digested genomic DNA. (B) Signal intensities were analysed using the Dot blot analyser from ImageJ software to generate plots. A.U. Arbitrary Units.

A6.3 Telomere mutation screening in the WI38V13/2RA shCtIPA1 clone by TVR mapping

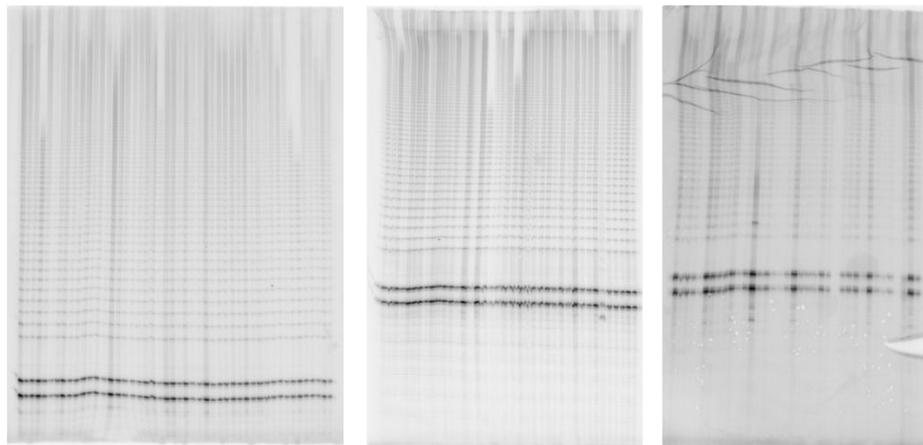


Figure A.9 Telomere mutation screening by TVR mapping in the WI38V13/2RA clone shCtIPA1. TVR was conducted in 122 STELA products from the shCtIPA1 using a specific primer for the TTAGGG repeats (TAG-TEL X). No mutations in the interspersion of telomere repeats were identified in the molecules analysed. All the STELA products analysed corresponded to the map 1 (see Figure 5.11) of the progenitor WI38V13/2RA cell line.

A6.4 *CtIP* expression analysis by qPCR after transfection with siRNA against *CtIP*

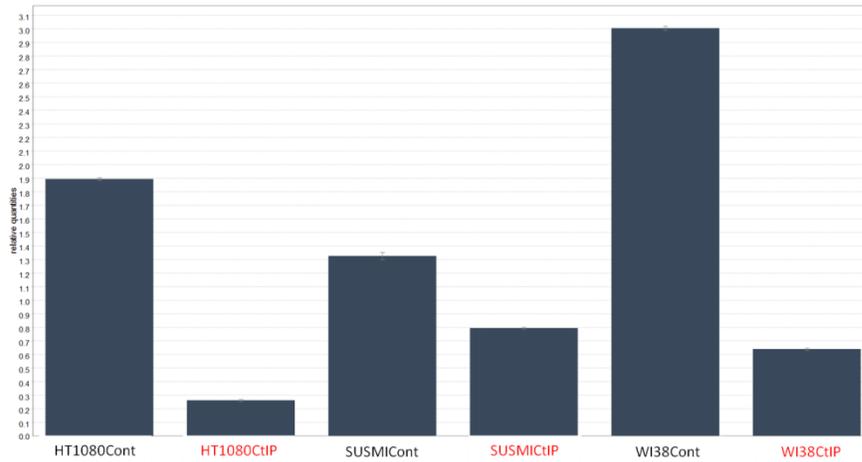


Figure A.10 Technical replica of qPCR analysis of *CtIP* expression in HT1080, SUSM1 and WI38V13/2RA 48hrs after transfection with siCtIP or siControl. The downregulation achieved was 85%, 40% and 80% respectively. Data was normalised using GAPDH, β -Actin and UBC as reference genes. The reactions were done in triplicate and data analysed using the qbase^{plus} software (Biogazelle).

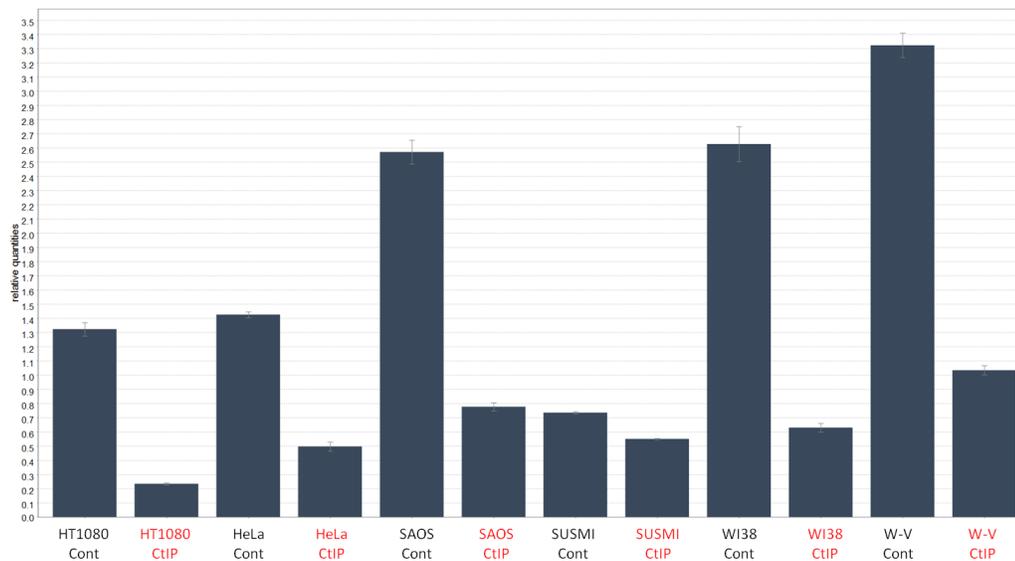


Figure A.11 Technical replicates of the qPCR analysis of *CtIP* expression in the HT1080, HeLa, SAOS, SUSMI, WI38V13/2RA and W-V 48hrs after transfection with siCtIP or siControl. Data was normalised using GAPDH, β -Actin and UBC as reference genes. The downregulation achieved was 85%, 75%, 80%, 50%, 75% and 70% respectively. Reactions were done in triplicate. Data was analysed using qbase^{plus} software (Biogazelle).

A6.5 C-circle quantification in the ALT cell lines after transfection with siRNA against *CtIP*

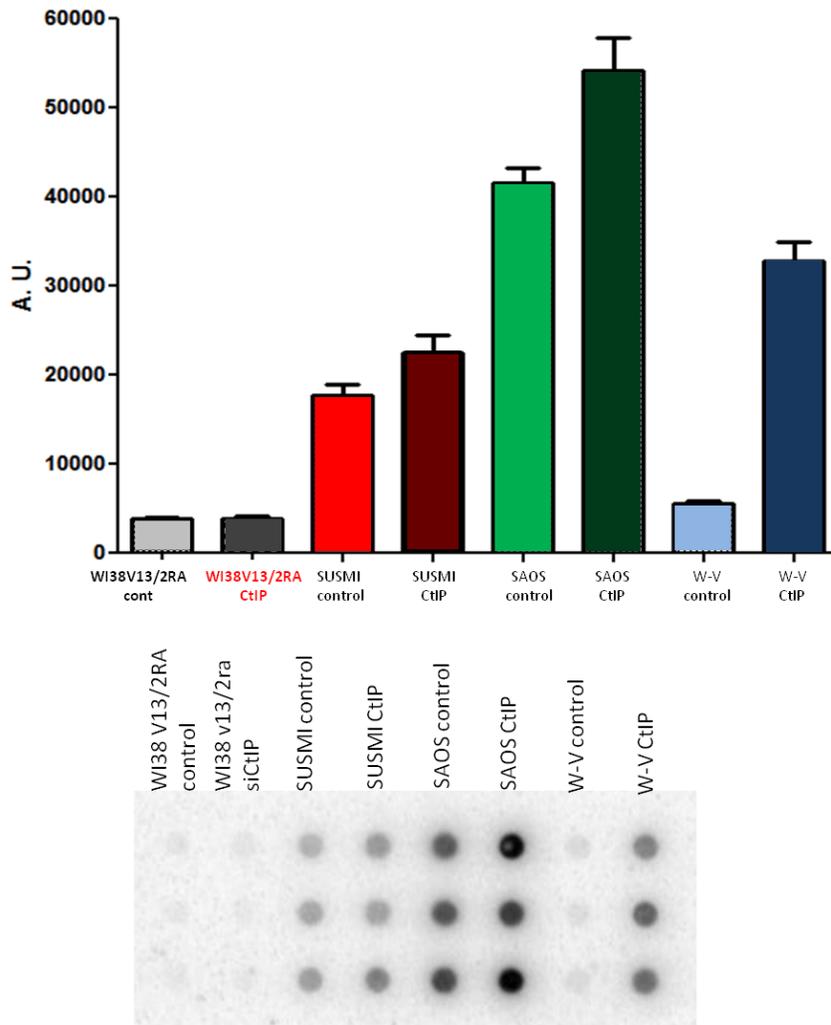


Figure A.12 Replica of the quantification of C-circle DNA in WI38V13/2RA, SUSM1, SAOS and W-V ALT+ cell lines 48hrs after the second transfection with siCtIP or siControl. The C-circle abundance did not change in the WI38V13/2RA cell line, while in the rest of the cell lines there was an increase on the level of C-circle DNA. Dot blot after hybridisation with a ³²P radiolabelled telomere probe. Signal intensities were measured using dot blot analyser from ImageJ software. Data was analysed using GraphPad Prism 5 software. A.U. Arbitrary Units.

APPENDIX 7 Analysis of samples with CIHHV-6

A7.1 Southern blots of STELAs in blood DNA from HHV-6 positive siblings used for telomere length analysis on telomeres XpYp, 12q, 17p and the HHV-6 associated telomere.

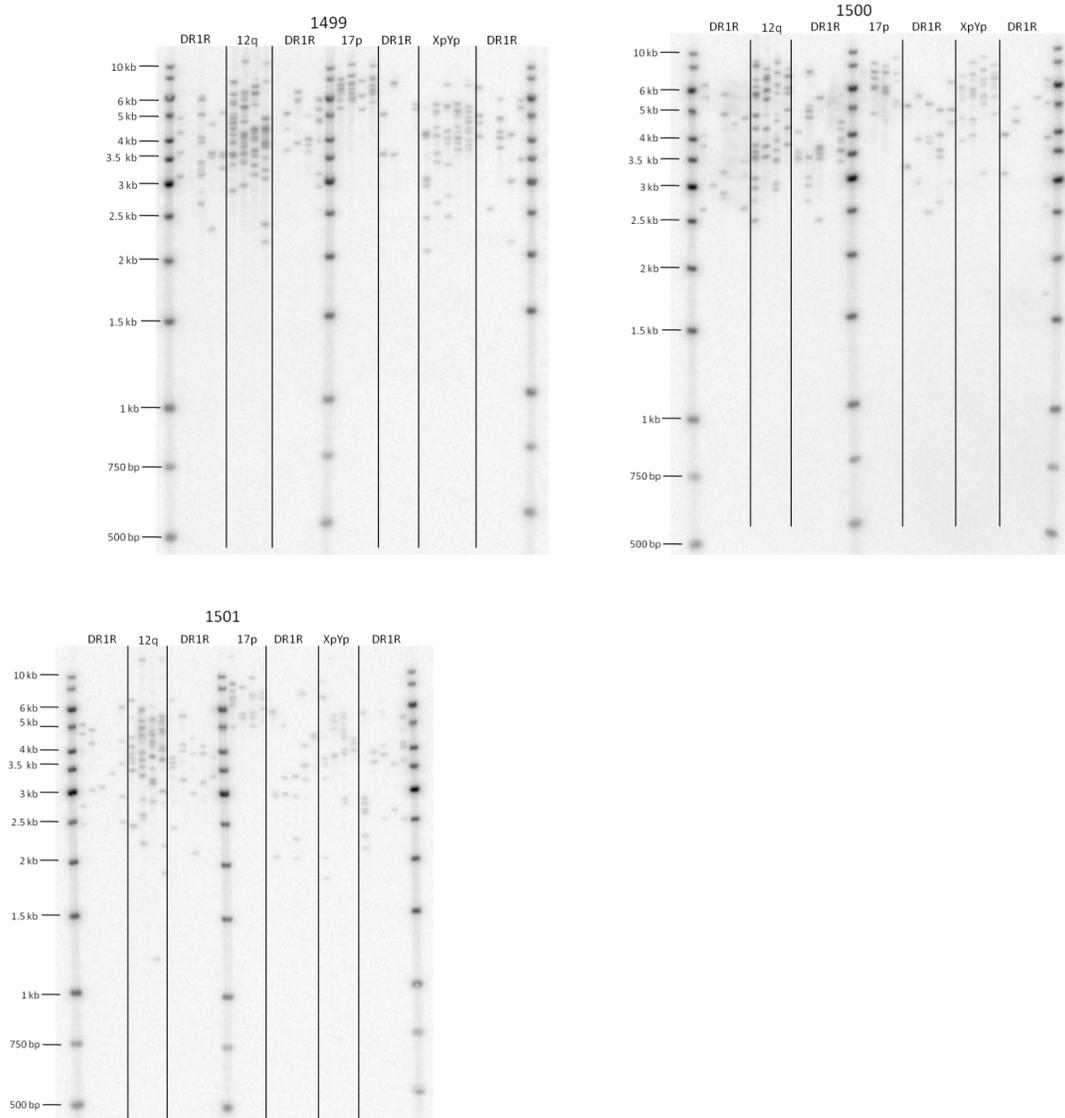
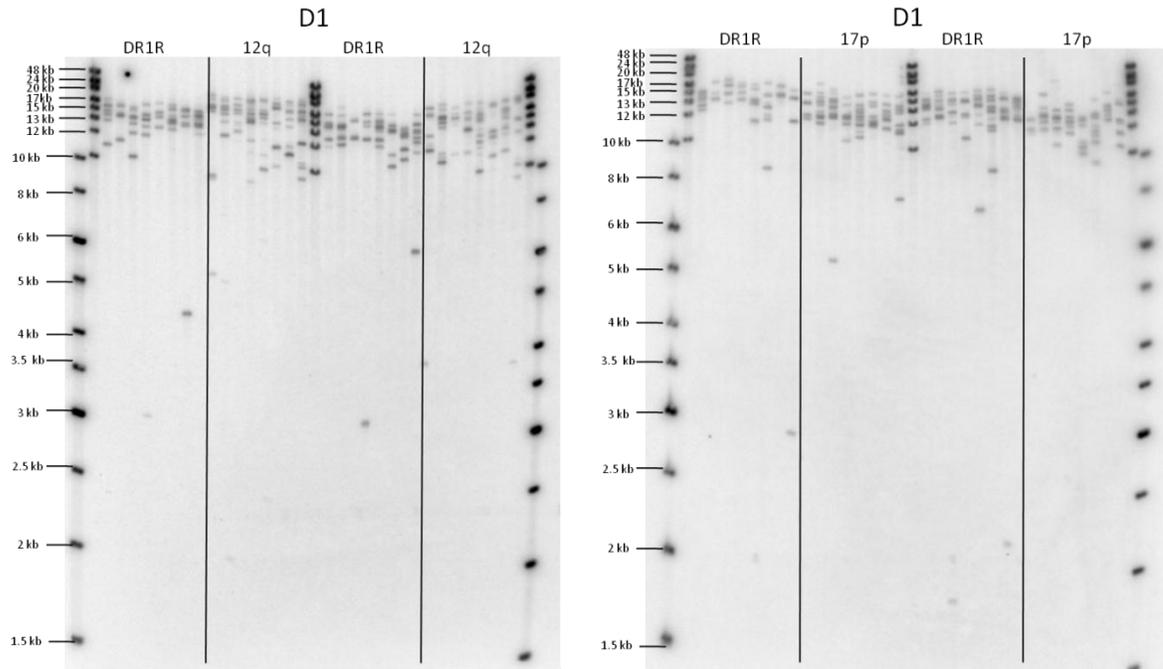
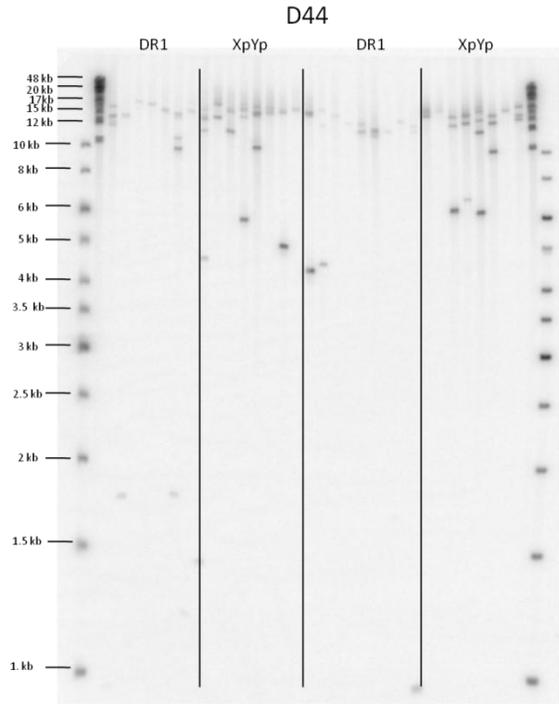
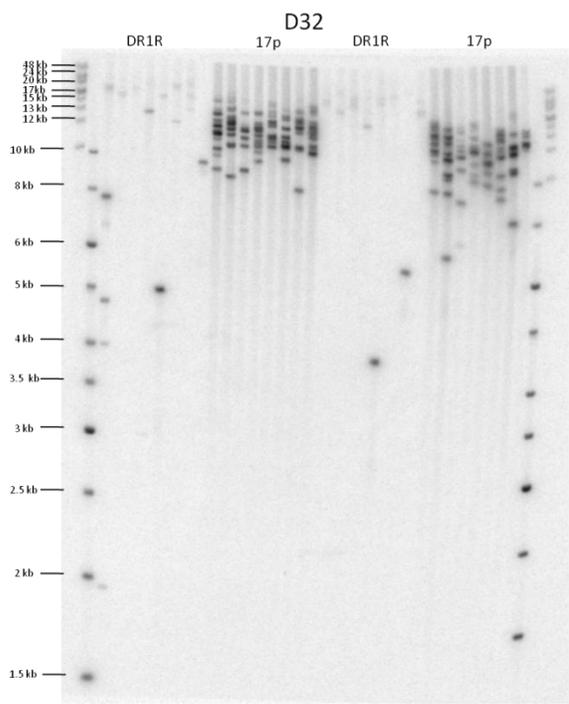
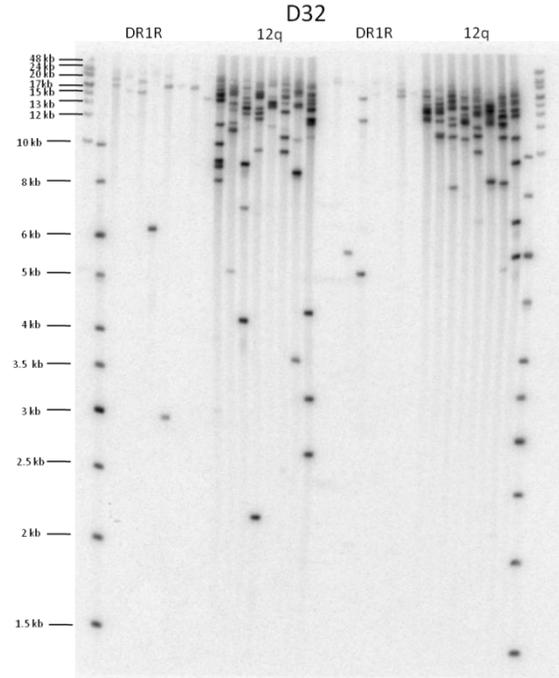
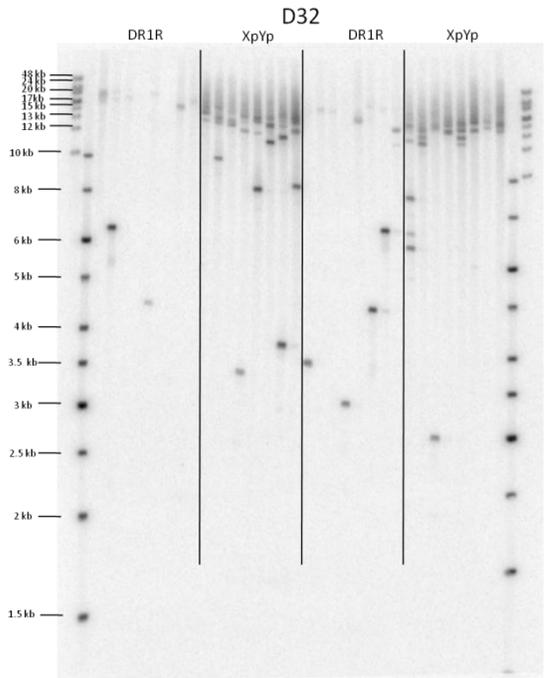
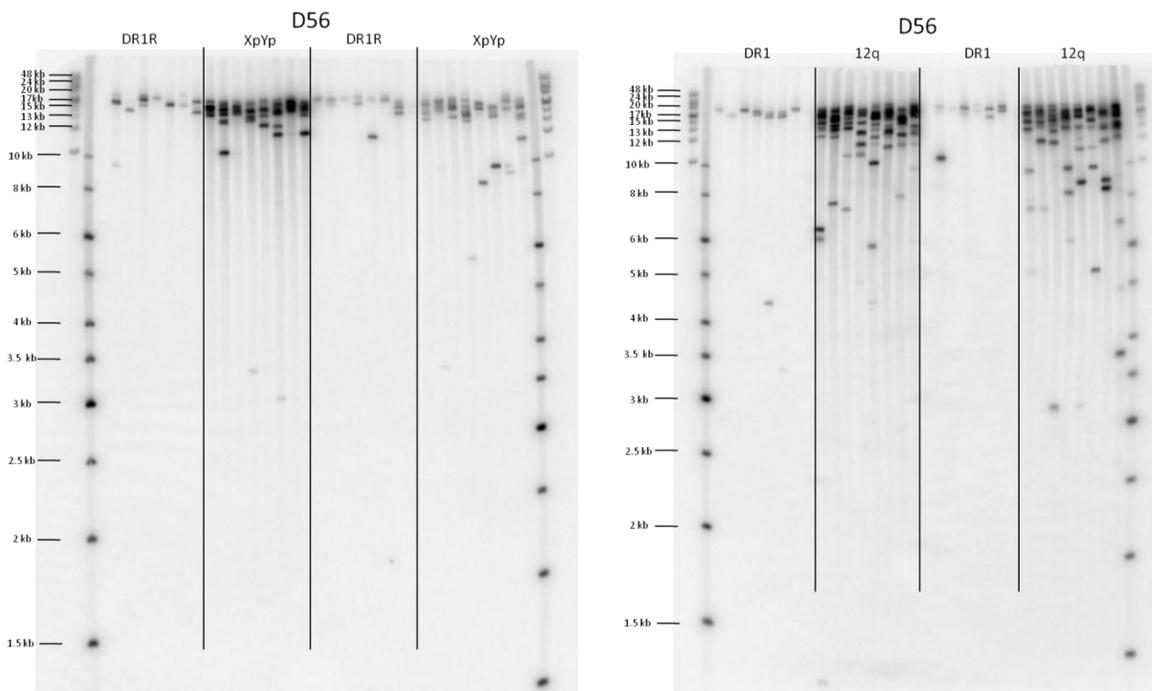
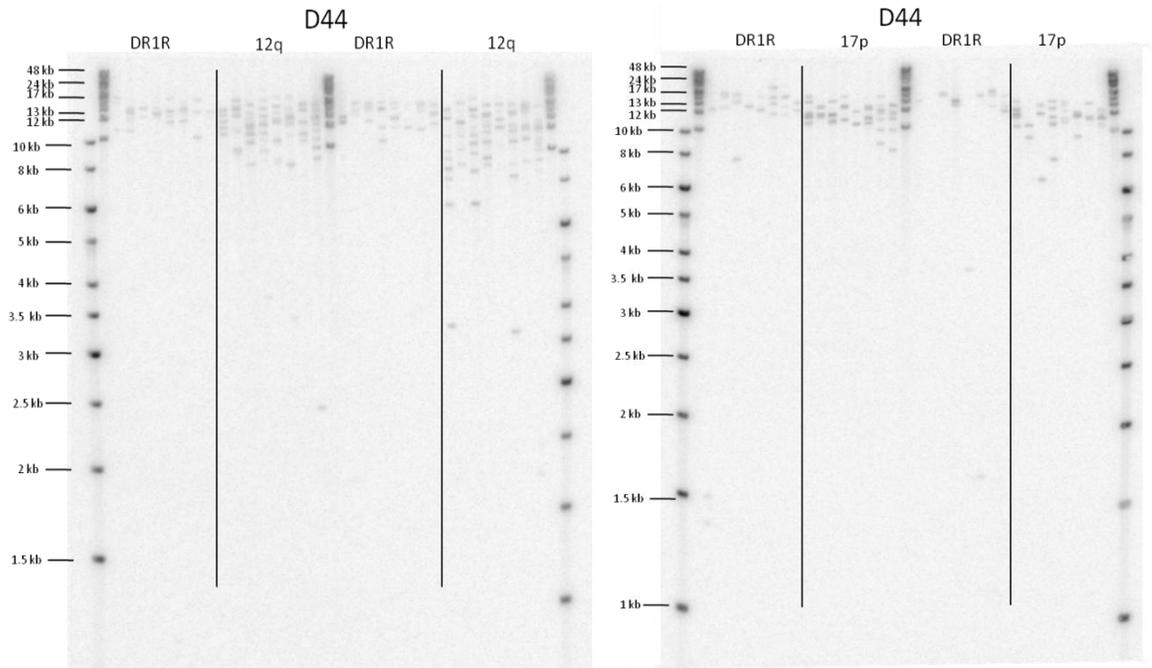


Figure A.13 STELA Southern blot analysis in blood DNA from the HHV-6 positive siblings (1499-1501) used for telomere length analysis. STELA was conducted using the HHV-6 (DR1R), XpYp, 12q and 17p telomere primers with Teltail and Telorette2. Genomic DNA was digested with *EcoRI* and 250pg of *EcoRI* digested DNA were used per reaction. The products were detected following agarose gel size fractionation by Southern blot hybridisation to a ^{32}P radiolabelled telomere repeat probe.

A7.2 Southern blots of STELAs in sperm DNA from HHV-6 positive semen donors used for telomere length analysis on telomeres XpYp, 12q, 17p and the HHV-6 associated telomere.







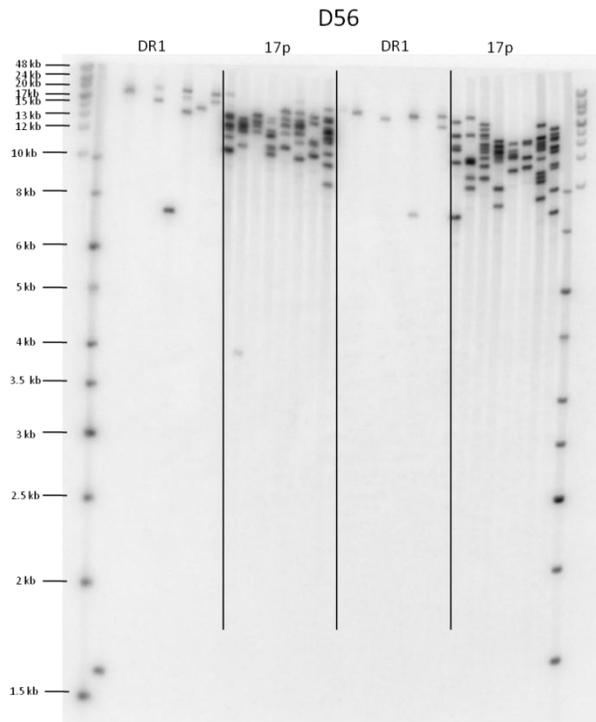


Figure A.14 STELA Southern blot analysis in sperm DNA from the HHV-6 positive donors used for telomere length analysis. STELA was conducted using the HHV-6 (DR1R), XpYp, 12q and 17p telomere primers with Teltail and Telorette2. Genomic DNA was digested with *EcoRI* and 300 pg of digested DNA were used per reaction. The products were detected following agarose gel size fractionation by Southern blot hybridisation to a ^{32}P radiolabelled telomere repeat probe.

A7.3 Histograms of the data used for the telomere length analysis in the four sperm donors

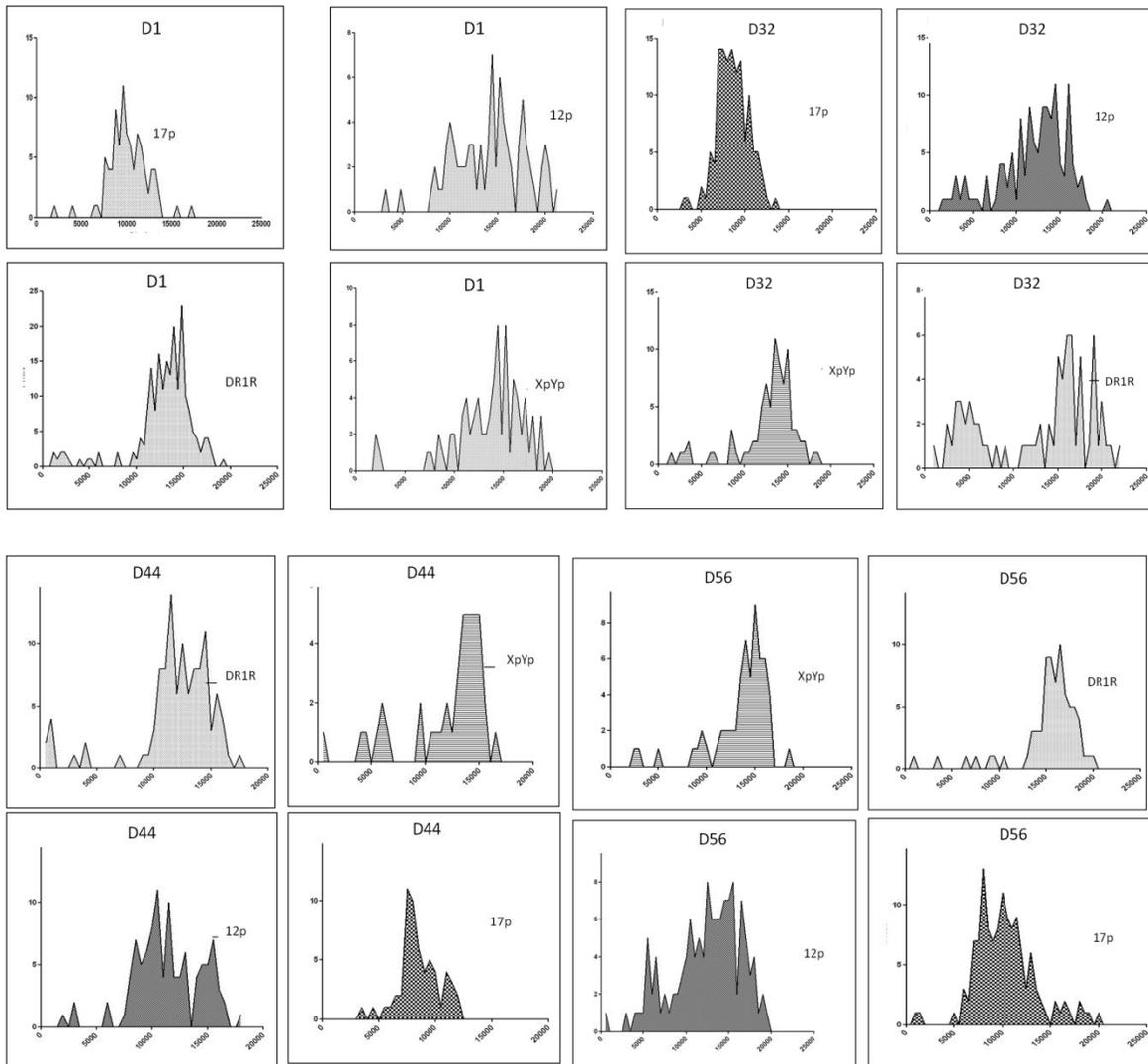
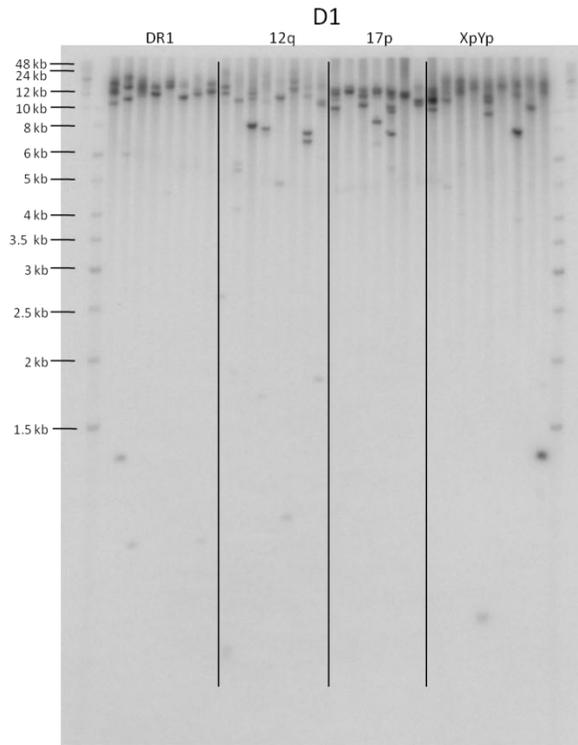
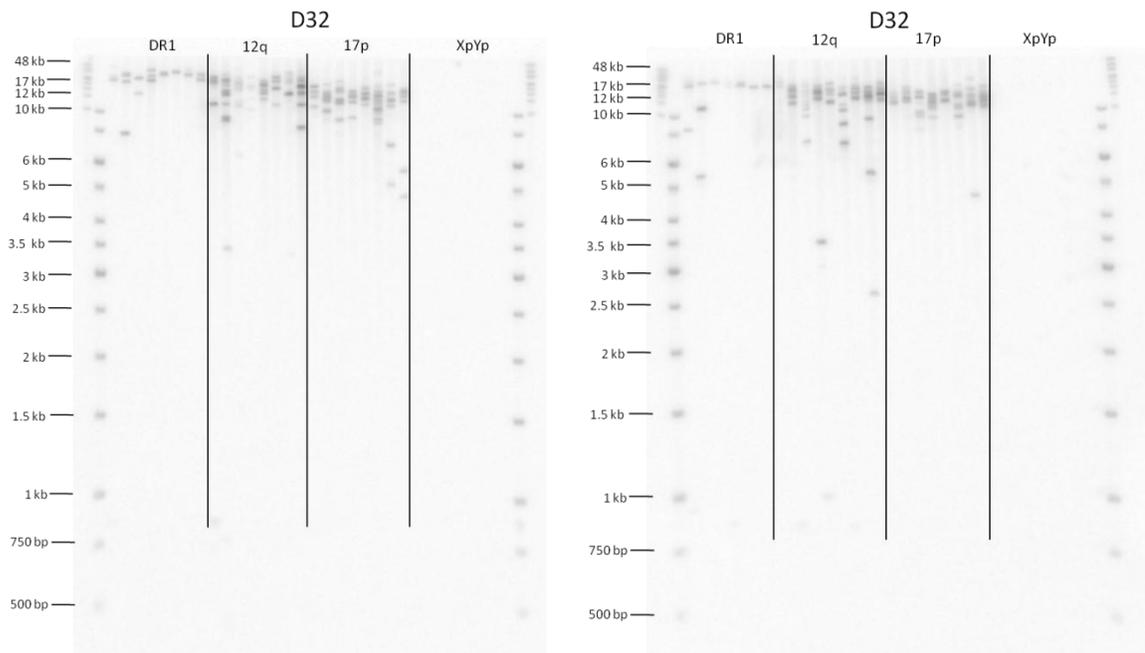
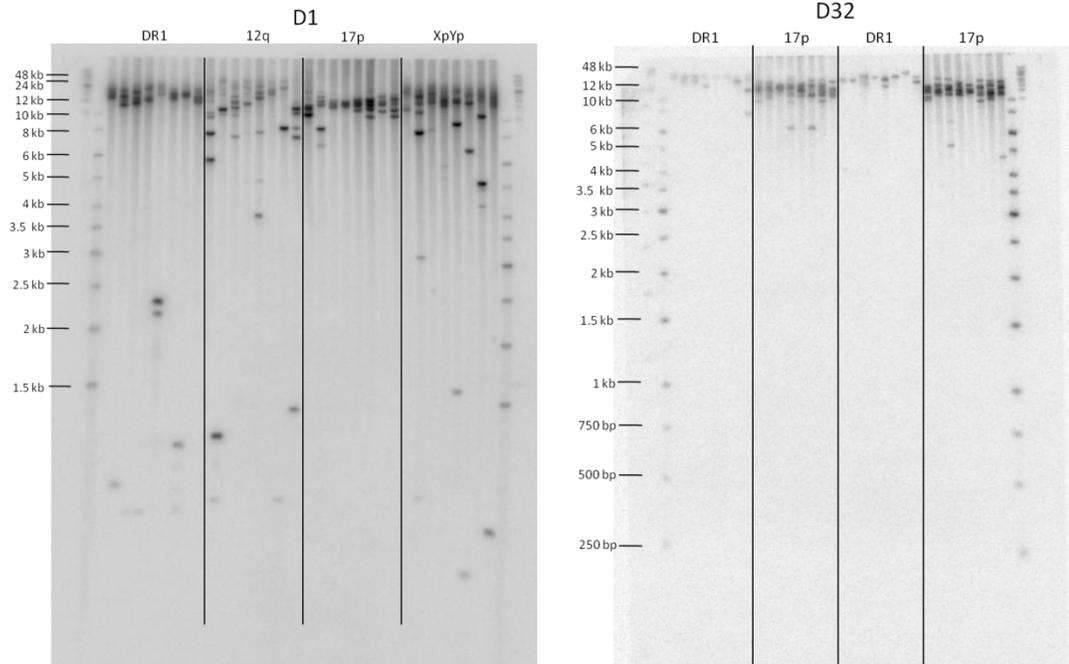
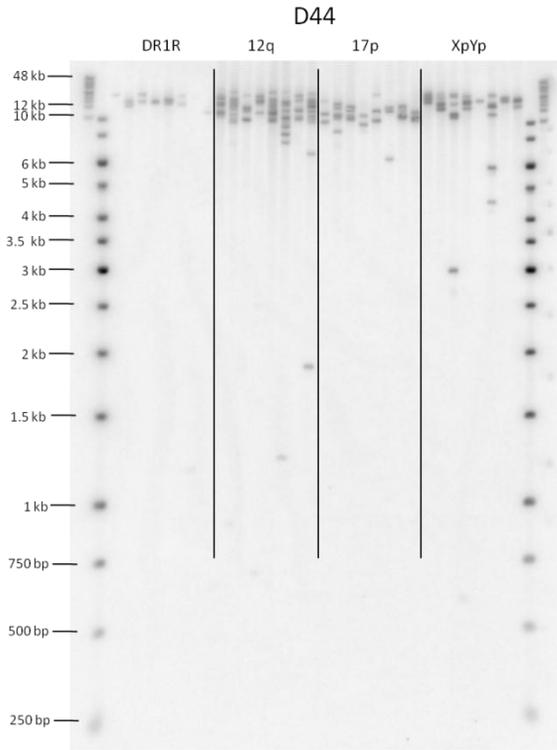
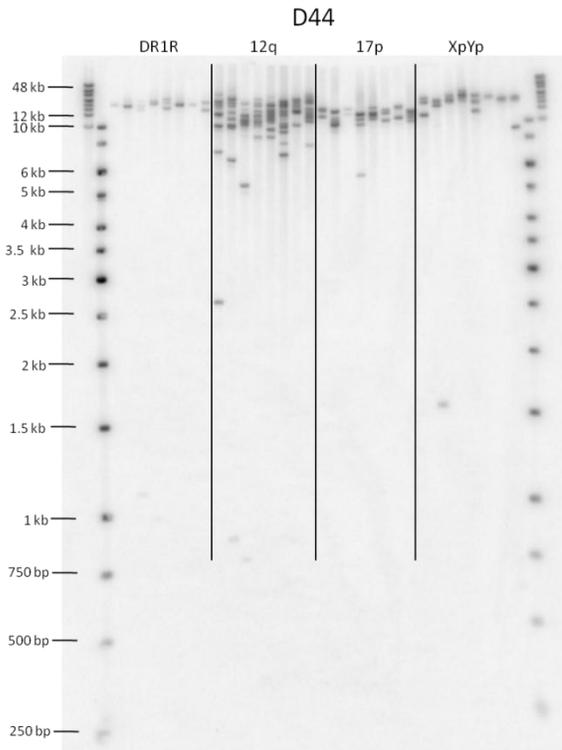
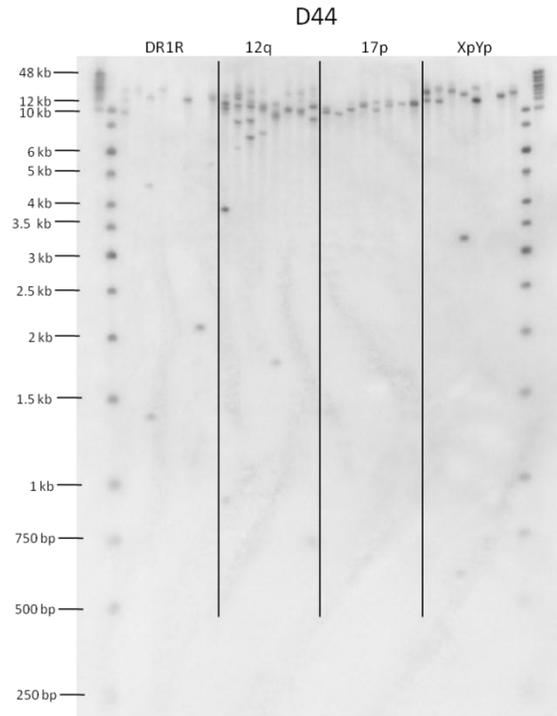
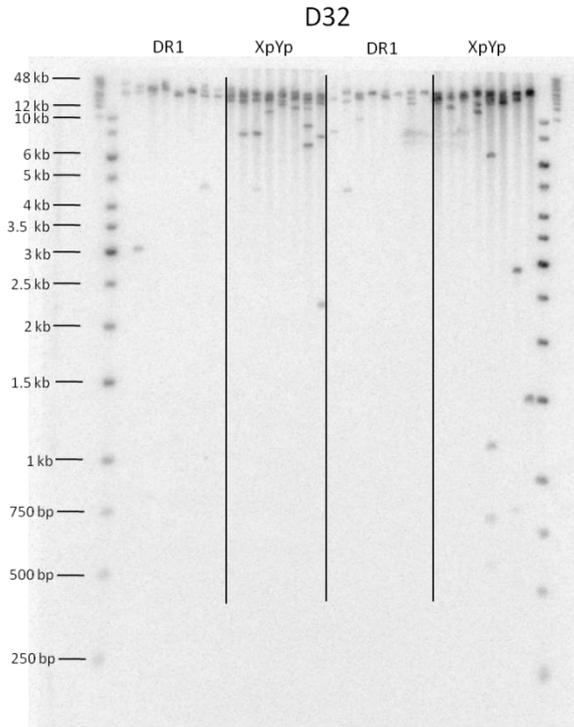


Figure A.15 Histograms generated from the data used in the telomere length analysis from the four sperm donors. Distribution of the data resembles normal distribution in most of the samples except for the HHV-6 associated telomere (DR1R) in donor 32 where it looks like there are two populations of molecules.

A7.4 Southern blots of STELAs in sperm DNA from HHV-6 positive semen donors used for analysis of short telomeres on the XpYp, 12q, 17p and HHV-6 associated telomere.







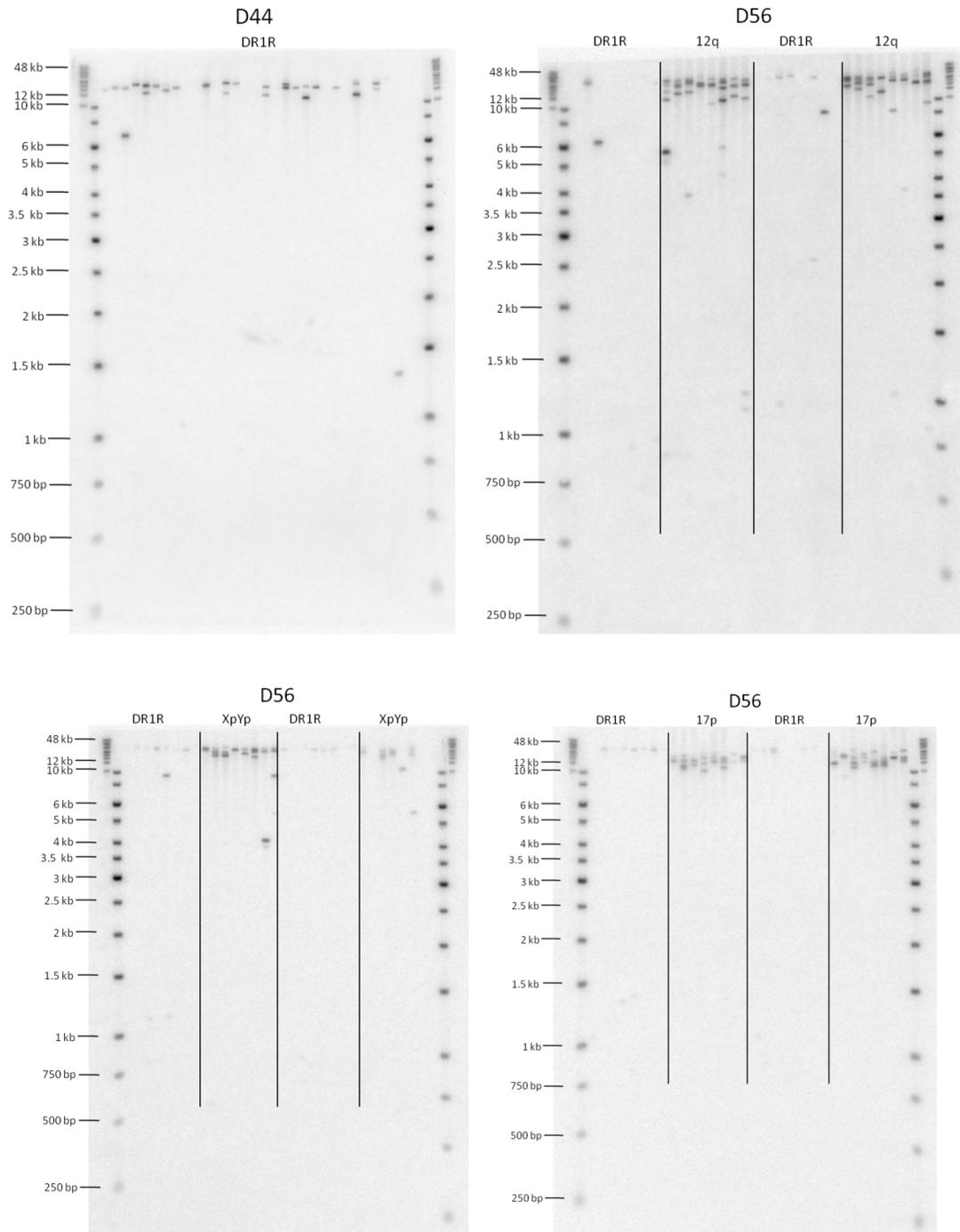


Figure A.16 STELA Southern blot analysis in sperm DNA from the HHV-6 positive donors used for analysis of short telomeres. STELA was conducted using the HHV-6 (DR1R), XpYp, 12q and 17p telomere primers with Teltail and Telorette2. Genomic DNA was digested with *EcoRI* and 300 pg of digested DNA were used per reaction. The products were detected following agarose gel size fractionation by Southern blot hybridisation to a ^{32}P radiolabelled telomere repeat probe.

A7.6.2 Sequence from sperm donor 1 harbouring CIHV-6A

ACTTGTAGGCTTAGGATGTAATGTAGACATTCTTGCTCGGGCCAAGGTAAAGGAATGCTAgACCTTTGTGC
 AGACGCAGGTAGGAATGAACATGACGAAGGGAATGTGACCCAAGAGTAGCCACCAATAGTTGGAGAGC
 GGAAGACGATATGGCGCGTGCCCGCCACGCACCGTTAGCGGCATCCTAGAGTAAGAATAGAGGGGGCGGA
 CAAGTGTAGAAACAAGCGGTGCGAGGTGGAGGAAAAAGGTGAAGTTGAAGCGGCAGGGGGTGATGGTG
 TGAAAGAGGGGCAGCGTTAGGTGCGGATGCGAGGATTAGTGTAGGGTTGTGGACAGAGTTATGCGGTTA
 TGGCTAGGGTTATGGCTAGGGTTATGGCTAGGGTTATGCTAGGGTTATGGCTAGGGTTATTGCTAGGGTTA
 TTGTTAGGGTTATTGTTAGgGTTATTGTTAGGGTTATTGCTAGGGTTATGGCTATAGTTATGGCTAGGGTTA
 GTRACTGTGGCTAGGGTTAGGGTTAGTACTTTGGCTGGGGCTAGGCTTAGTGTAGGGATAGTTTTATGGTT
 AGTGTATGGTTGGTGTGTTAGGGTTAGTGTTCGGTTAGGGCTAGTACTTTGGTTACGGCTAGAGTTTTCG
 GGTTAGGGTTAGGGTTAGGGTTAGGGTTA**TAGGGTAGGGCTAGGCCTAGGGTTAGGGTTAGGGTTAGGGTTAGG**
GCTAGGCCTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGTAGGGTTAGGGTTAGGGCTAGGGCT
AGGGTTAGGGTTAGGGTTAGGGCTAGATCTAGGGTTAGGGTTAGGGTTAGGGCTAGGCCTAGGGTTAG
GGTTAGGGTTAGGGTTAGGGTTAGGGCTAGACGTAGGGTTAGGGTTAGGGATAGACCTAGGGTTAGGG
TTAGGGTTAGGGCTAGACCTAGGGTTAGGGTTAGGGTTAGGGTTAGGGCTAGACCTAGGGTTAGGGTTA
GGGTTAGGGTTAGGGTTAGGGCTAGACCTAGGGTTAGGGTTAGGGTTAGGGCTAGACCTAGGGTTAGG
GTTAGGGTTAGGGCTAGATCTAGGGTTAGGGTTAGGGTTAGGGCTAGACCTAGGGTTAGGGTTAGGGTT
AGGGCTAGACCTAGGGTTAGGGTTAGGGTTAGGGCTAGACCTAGGGTTAGGGTTAGGGTTAGGGCTAG
ACCTAGGGTTAGGGTTAGGGTTAGGAGTAGACGTAGGGTTAGGGTTAGGGATAGACCTAGGGTTAGGG
TTAGGGTTAGGGCTAGACCTAGGGTTAGGGTTAGGGTTAGGGTTAGGGATAGACCTAGGGTTAGGGTTA
GGGCTAGACCTAGGGTTAGGGTTAGGGCTAGACCTAGGGTTAGGGTTAGGGCTAGATCTAGGGTTAG
GGTTAGGGCTAGACCTAGGGTTAGGGTTAGGGTTAGGGCTAGACCTAGGGTTAGGGTTAGGGTTAGGG
CTAGATCTAGGGTTAGGGTTAGGGTTAGGGCTAGACCTAGGGTTAGGGTTAGGGTTAGGGCTAGACCTA
GGGTTAGGGTTAGGGTTAGGGCTAGACCTAGGGTTAGG

Figure A.19 Sequence of the distal chromosome-HHV-6A junction obtained from STELA products from sperm DNA from donor 1 harbouring CIHV-6A. The first base at the 5' end represents the base 687 from the DR_L from the HHV-6A strain U1102 (GenBank: X83413.1). In bold the T1 region of imperfect telomeric repeats containing TTAGGG repeats (black) and variant repeats (in blue) from the HHV-6 genome. The 3' end of the molecule does not finish in a pure array of TTAGGG repeats. The Pac1 element is missing.

A7.6.3 Sequence from sperm donor 56 harbouring CIHHV-6B

AGTGTAGGGCTAGGGTTAGGGCTATGGTTAGTGCTATGGTTAGGGCTATGGTTAGGGTTAGGGTTAGG
GCTATGGTTAGGGCTATGGTTAGGGCTATGGTTAGGGCTAGGGTTAGGGCTAGGGTTAGGGCTAGGGTT
AGGGCTAGGGTTAGGGCTAGGGTTAGGGCTAGGGCTAGGGCTAGGGCTAGGGCTAGGGCTAGGGCTAGG
GGCTAGGGCTAGGGCTAGGGCTAGGGTTAGGGCTAGGGCTAAGGCTAGGGCTAGGGCTAGGGCTAGG
GCTAGGGTTAGGGCTTTGACTAGGGCCAGACTTAGTGTTAGTGTTAGGGTTATGGCTAAGGCTAAGGAT
AGACTTACTGTTAGTGTTAGGGTTAGGGCTAGGGCTAAGGATAGACTTACTGTTAGTGTTAGGGTTAGG
GCT//2KB//
TTAGGGTTAGAGTTAGACCTAGGGTTAGGGTTAGAGTTAGACCTAGGGTTAGGGTTAGGGTTAGGGTTA
GGGTTAGGGTTAGGGTTAGGGTTAGACCTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGG
GTTAGGGTTGGAGTTAGACCTAGGGTTAGGGTTAGGGTTGGAGTTAGACCTAGGGTTAGGGTTAGGGTT
AGAGTTAGACCTAGGGTTAGGGTTAGACCTAGGGTTAGGGTTAGGGTTGGAGTTAGACCTAGGGTTAGG
GTTAGGGTTAGAGTTAGACCTAGGGTTAGGGTTAGGGTTAGGGTTAGACCTAGGGTTAGGGTTAGGGTT
AGACCTAGGGTTAGGGTTAGACCTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAG
GGTTGGAGTTAGACCTAGGGTTAGGG

Figure A.20 Sequence of the distal chromosome-HHV-6A junction obtained from STELA products from sperm DNA from donor 56 harbouring CIHHV-6B. In bold the T1 region of imperfect telomeric repeats containing TTAGGG repeats (black) and variant repeats (in blue) from the HHV-6 genome. The molecule does not finish in a pure array of TTAGGG repeats. The Pac1 element is missing. There is a gap of ~1.5kb that could not be sequenced because the repetitive nature of the molecule.

BIBLIOGRAPHY

Ablashi D, Agut H and Berneman Z, 1993. Human herpesvirus-6 strain groups: a nomenclature. - *Arch Virol.*1993;129(1-4):363-6., .

Ablashi DV, Salahuddin SZ, Josephs SF, Imam F, Lusso P, Gallo RC, Hung C, Lemp J, Markham, P.D. 1987. HBLV (or HHV-6) in human cell lines. - *Nature.*1987 Sep 17-23;329(6136):207., .

Achour A, Malet I, Deback C, Bonnafous P, Boutolleau D, Gautheret-Dejean A, Agut H, 2009. Length variability of telomeric repeat sequences of human herpesvirus 6 DNA. - *J Virol Methods.*2009 Jul;159(1):127-30.Epub 2009 Mar 14., .

Ahnesorg P, Smith P and Jackson SP, 2006. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. - *Cell.*2006 Jan 27;124(2):301-13., .

Al-Buhtori M, Moore L, Benbow EW, Cooper RJ, 2011. Viral detection in hydrops fetalis, spontaneous abortion, and unexplained fetal death in utero. - *J Med Virol.*2011 Apr;83(4):679-84.doi: 10.1002/jmv.22007., .

Andersen SD, Keijzers G, Rampakakis E, Engels K, Luhn P, El-Shemerly M, Nielsen FC, Du Y, May A, Bohr VA, Ferrari S, Zannis-Hadjopoulos M, Fu H, FAU and Rasmussen LJ, 2012. 14-3-3 checkpoint regulatory proteins interact specifically with DNA repair protein human exonuclease 1 (hEXO1) via a semi-conserved motif. - *DNA Repair (Amst).*2012 Mar 1;11(3):267-77.Epub 2012 Jan 4., .

Andrew T, Aviv A, Falchi M, Surdulescu GL, Gardner JP, Lu X, Kimura M, Kato BS, Valdes AM, and Spector TD, 2006. Mapping genetic loci that determine leukocyte telomere length in a large sample of unselected female sibling pairs. - *Am J Hum Genet.*2006 Mar;78(3):480-6.Epub 2006 Jan 6., .

Arbuckle JH, Medveczky MM, Luka J, Hadley SH, Luegmayer A, Ablashi D, Lund TC, Tolar J, De Meirleir K, Montoya JG, Komaroff AL, Ambros PF, and Medveczky PG, 2010. The latent human herpesvirus-6A genome specifically integrates in telomeres of human chromosomes in vivo and in vitro. - *Proc Natl Acad Sci U S A.*2010 Mar 23;107(12):5563-8.Epub 2010 Mar 8., .

Arbuckle JH, and Medveczky PG, 2011. The molecular biology of human herpesvirus-6 latency and telomere integration. - *Microbes Infect.*2011 Aug;13(8-9):731-41.Epub 2011 Mar 31., .

Ashshi AM, Cooper RJ, Klapper PE, Al-Jiffri O, and Moore L, 2000. Detection of human

herpes virus 6 DNA in fetal hydrops. - *Lancet*.2000 Apr 29;355(9214):1519-20., .

Aubin JT, Collandre H, Candotti D, Ingrand D, Rouzioux C, Burgard M, Richard S, Huraux JM, and Agut H, 1991. Several groups among human herpesvirus 6 strains can be distinguished by Southern blotting and polymerase chain reaction. - *J Clin Microbiol*.1991 Feb;29(2):367-72., .

Bailey SM, Cornforth MN, Kurimasa A, Chen DJ, and Goodwin EH, 2001. Strand-specific postreplicative processing of mammalian telomeres. - *Science*.2001 Sep 28;293(5539):2462-5., .

Baird DM, Britt-Compton B, Rowson J, Amso NN, Gregory L, and Kipling D, 2006. Telomere instability in the male germline. - *Hum Mol Genet*.2006 Jan 1;15(1):45-51.Epub 2005 Nov 25., .

Baird DM, Coleman J, Rosser ZH, and Royle NJ, 2000. High levels of sequence polymorphism and linkage disequilibrium at the telomere of 12q: implications for telomere biology and human evolution. - *Am J Hum Genet*.2000 Jan;66(1):235-50., .

Baird DM, Davis T, Rowson J, Jones CJ, and Kipling D, 2004. Normal telomere erosion rates at the single cell level in Werner syndrome fibroblast cells. - *Hum Mol Genet*.2004 Jul 15;13(14):1515-24.Epub 2004 May 18., .

Baird DM, Jeffreys AJ, and Royle NJ, 1995. Mechanisms underlying telomere repeat turnover, revealed by hypervariable variant repeat distribution patterns in the human Xp/Yp telomere. - *EMBO J*.1995 Nov 1;14(21):5433-43., .

Baird DM, Rowson J, Wynford-Thomas D, and Kipling D, 2003. Extensive allelic variation and ultrashort telomeres in senescent human cells. - *Nat Genet*.2003 Feb;33(2):203-7.Epub 2003 Jan 21., .

Bates M, Monze M, Bima H, Kapambwe M, Clark, D, Kasolo FC, and Gompels UA, 2009. Predominant human herpesvirus 6 variant A infant infections in an HIV-1 endemic region of Sub-Saharan Africa. - *J Med Virol*.2009 May;81(5):779-89., .

Baynton K, Otterlei M, Bjoras M, von Kobbe C, Bohr VA, and Seeberg E, 2003. WRN interacts physically and functionally with the recombination mediator protein RAD52. - *J Biol Chem*.2003 Sep 19;278(38):36476-86.Epub 2003 May 15., .

Bhattacharyya S, Keirse J, Russell B, Kavecansky J, Lillard-Wetherell K, Tahmaseb K, Turchi JJ, and Groden J, 2009. Telomerase-associated protein 1, HSP90, and topoisomerase IIalpha associate directly with the BLM helicase in immortalized cells using ALT and modulate its helicase activity using telomeric DNA substrates. - *J Biol Chem*.2009

May 29;284(22):14966-77.Epub 2009 Mar 27., .

Black JB, and Pellett PE, 1999. Human herpesvirus 7. - *Rev Med Virol.*1999 Oct-Dec;9(4):245-62., .

Blackburn EH, and Gall JG, 1978. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. - *J Mol Biol.*1978 Mar 25;120(1):33-53., .

Blackford AN, Schwab RA, Nieminuszczy J, Deans AJ, West SC and Niedzwiedz W, 2012. The DNA translocase activity of FANCM protects stalled replication forks. - *Hum Mol Genet.*2012 Feb 8., .

Bohr VA, 2008. Rising from the RecQ-age: the role of human RecQ helicases in genome maintenance. - *Trends Biochem Sci.*2008 Dec;33(12):609-20.Epub 2008 Oct 14., .

Bolderson E, Richard DJ, Edelman W, and Khanna KK, 2009. Involvement of Exo1b in DNA damage-induced apoptosis. - *Nucleic Acids Res.*2009 Jun;37(10):3452-63.Epub 2009 Apr 1., .

Bonetti D, Martina M, Clerici M, Lucchini G, and Longhese MP, 2009. Multiple pathways regulate 3' overhang generation at *S. cerevisiae* telomeres. - *Mol Cell.*2009 Jul 10;35(1):70-81., .

Bradford MM, 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. - *Anal Biochem.*1976 May 7;72:248-54., .

Brock GJ, Charlton J, and Bird A, 1999. Densely methylated sequences that are preferentially localized at telomere-proximal regions of human chromosomes. - *Gene.*1999 Nov 29;240(2):269-77., .

Bryan TM, Englezou A, Dalla-Pozza L, Dunham MA, and Reddel RR, 1997. Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. - *Nat Med.*1997 Nov;3(11):1271-4., .

Bugreev, D.V., Yu, X., Egelman, E.H. and Mazin, A.V., 2007. Novel pro- and anti-recombination activities of the Bloom's syndrome helicase. *Genes & Development*, **21**3085-3094.

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, and Wittwer CT, 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. - *Clin*

*Chem.*2009 Apr;*55(4):611-22.Epub 2009 Feb 26., .*

Calnek BW, 2001. Pathogenesis of Marek's disease virus infection. - *Curr Top Microbiol Immunol.*2001;*255:25-55., .*

Caselli E, Bracci A, Galvan M, Boni M, Rotola A, Bergamini C, Cermelli C, Dal Monte P, Gompels UA, Cassai E, and Di Luca D, 2006. Human herpesvirus 6 (HHV-6) U94/REP protein inhibits betaherpesvirus replication. - *Virology.*2006 Mar 15;*346(2):402-14.Epub 2005 Dec 20., .*

Caselli E, and Di Luca D, 2007. Molecular biology and clinical associations of Roseoloviruses human herpesvirus 6 and human herpesvirus 7. - *New Microbiol.*2007 Jul;*30(3):173-87., .*

Catusse J, Clark DJ, and Gompels UA, 2009. CCR5 signalling, but not DARC or D6 regulatory, chemokine receptors are targeted by herpesvirus U83A chemokine which delays receptor internalisation via diversion to a caveolin-linked pathway. - *J Inflamm (Lond).*2009 Jul 30;*6:22., .*

Cerone MA, Autexier C, Londono-Vallejo JA, and Bacchetti S, 2005. A human cell line that maintains telomeres in the absence of telomerase and of key markers of ALT. - *Oncogene.*2005 Nov 24;*24(53):7893-901., .*

Cesare AJ, Kaul Z, Cohen SB, Napier CE, Pickett HA, Neumann AA, and Reddel RR, 2009. Spontaneous occurrence of telomeric DNA damage response in the absence of chromosome fusions. - *Nat Struct Mol Biol.*2009 Dec;*16(12):1244-51.Epub 2009 Nov 22., .*

Cesare AJ, and Reddel RR, 2010. Alternative lengthening of telomeres: models, mechanisms and implications. - *Nat Rev Genet.*2010 May;*11(5):319-30.Epub 2010 Mar 30., .*

Chang S, Multani AS, Cabrera NG, Naylor ML, Laud P, Lombard D, Pathak S, Guarente L, and DePinho RA, 2004. Essential role of limiting telomeres in the pathogenesis of Werner syndrome. - *Nat Genet.*2004 Aug;*36(8):877-82.Epub 2004 Jul 4., .*

Chappell C, Hanakahi LA, Karimi-Busheri F, Weinfeld M, and West SC, 2002. Involvement of human polynucleotide kinase in double-strand break repair by non-homologous end joining. - *EMBO J.*2002 Jun 3;*21(11):2827-32., .*

Chatterjee K, Dandara C, Gyllensten U, van der Merwe L, Galal U, Hoffman M, and Williamson AL, 2010. A Fas gene polymorphism influences herpes simplex virus type 2 infection in South African women. - *J Med Virol.*2010 Dec;*82(12):2082-6., .*

Chen PL, Liu F, Cai S, Lin X, Li A, Chen Y, Gu B, Lee EY, and Lee WH, 2005. Inactivation of

CtIP leads to early embryonic lethality mediated by G1 restraint and to tumorigenesis by haploid insufficiency. - *Mol Cell Biol.*2005 May;25(9):3535-42., .

Chen, Q., Ijima, A. and Greider, C.W., 2001. Two survivor pathways that allow growth in the absence of telomerase are generated by distinct telomere recombination events. *Mol.Cell.Biol.*, 211819-1827.

Cheek CF, Wu L, Garcia PL, Janscak P, and Hickson ID, 2005. The Bloom's syndrome helicase promotes the annealing of complementary single-stranded DNA. - *Nucleic Acids Res.*2005 Jul 15;33(12):3932-41.Print 2005., .

Chinnadurai G, 2006. CtIP, a candidate tumor susceptibility gene is a team player with luminaries. - *Biochim Biophys Acta.*2006 Jan;1765(1):67-73.Epub 2005 Oct 6., .

Clark DA, Nacheva EP, Leong HN, Brazma D, Li YT, Tsao EH, Buyck HC, Atkinson CE, Lawson HM, Potter MN, and Griffiths PD, 2006. Transmission of integrated human herpesvirus 6 through stem cell transplantation: implications for laboratory diagnosis. - *J Infect Dis.*2006 Apr 1;193(7):912-6.Epub 2006 Feb 22., .

Compton SA, Choi JH, Cesare AJ, Ozgur S, and Griffith JD, 2007. Xrcc3 and Nbs1 are required for the production of extrachromosomal telomeric circles in human alternative lengthening of telomere cells. - *Cancer Res.*2007 Feb 15;67(4):1513-9., .

Counter CM, Avilion AA, LeFeuvre CE, Stewart NG, Greider CW, Harley CB, and Bacchetti S, 1992. Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. - *EMBO J.*1992 May;11(5):1921-9., .

Crabbe L, Jauch A, Naeger CM, Holtgreve-Grez H, and Karlseder J, 2007. Telomere dysfunction as a cause of genomic instability in Werner syndrome. - *Proc Natl Acad Sci U S A.*2007 Feb 13;104(7):2205-10.Epub 2007 Feb 6., .

Crabbe, L., Verdun, R.E., Haggblom, C.I. and Karlseder, J., 2004. Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science*, **306**1951-1953.

Crespan E, Czabany T, Maga G and Hubscher U, 2012. Microhomology-mediated DNA strand annealing and elongation by human DNA polymerases lambda and beta on normal and repetitive DNA sequences. - *Nucleic Acids Res.*2012 Feb 28., .

Daibata M, Taguchi T, Kamioka M, Kubonishi I, Taguchi H, and Miyoshi I, 1998. Identification of integrated human herpesvirus 6 DNA in early pre-B cell acute lymphoblastic leukemia. - *Leukemia.*1998 Jun;12(6):1002-4., .

Daibata M, Taguchi T, Nemoto Y, Taguchi H, and Miyoshi I, 1999. Inheritance of

chromosomally integrated human herpesvirus 6 DNA. - *Blood*.1999 Sep 1;94(5):1545-9., .

De Boeck G, Forsyth RG, Praet M, and Hogendoorn PC, 2009. Telomere-associated proteins: cross-talk between telomere maintenance and telomere-lengthening mechanisms. - *J Pathol*.2009 Feb;217(3):327-44., .

de Boer J, and Hoeijmakers JH, 2000. Nucleotide excision repair and human syndromes. - *Carcinogenesis*.2000 Mar;21(3):453-60., .

De Bolle L, Naesens L, and De Clercq E, 2005. Update on human herpesvirus 6 biology, clinical features, and therapy. - *Clin Microbiol Rev*.2005 Jan;18(1):217-45., .

de Lange T, 2005. Shelterin: the protein complex that shapes and safeguards human telomeres. - *Genes Dev*.2005 Sep 15;19(18):2100-10., .

Decottignies A, 2007. Microhomology-mediated end joining in fission yeast is repressed by pku70 and relies on genes involved in homologous recombination. - *Genetics*.2007 Jul;176(3):1403-15.Epub 2007 May 4., .

Denchi EL, and de Lange T, 2007. Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. - *Nature*.2007 Aug 30;448(7157):1068-71.Epub 2007 Aug 8., .

Dessain SK, Yu H, FAU Reddel RR, Beijersbergen RL, and Weinberg RA, 2000. Methylation of the human telomerase gene CpG island. - *Cancer Res*.2000 Feb 1;60(3):537-41., .

Devereux TR, Horikawa I, Anna CH, Annab LA, Afshari CA, and Barrett JC, 1999. DNA methylation analysis of the promoter region of the human telomerase reverse transcriptase (hTERT) gene. - *Cancer Res*.1999 Dec 15;59(24):6087-90., .

Dominguez G, Dambaugh TR, Stamey FR, Dewhurst S, Inoue N, and Pellett PE, 1999. Human herpesvirus 6B genome sequence: coding content and comparison with human herpesvirus 6A. - *J Virol*.1999 Oct;73(10):8040-52., .

Draskovic I, Arnoult N, Steiner V, Bacchetti S, Lomonte P, and Londono-Vallejo A, 2009. Probing PML body function in ALT cells reveals spatiotemporal requirements for telomere recombination. - *Proc Natl Acad Sci U S A*.2009 Sep 15;106(37):15726-31.Epub 2009 Aug 26., .

Du, X., Shen, J., Kugan, N., Furth, E.E., Lombard, D.B., Cheung, C., Pak, S., Luo, G., Pignolo, R.J., DePinho, R.A., Guarente, L. and Johnson, F.B., 2004. Telomere shortening exposes functions for the mouse Werner and Bloom syndrome genes. *Molecular & Cellular Biology*, **24**:8437-8446.

Dunham MA, Neumann AA, Fasching CL, and Reddel RR, 2000. Telomere maintenance by recombination in human cells. - *Nat Genet.*2000 Dec;26(4):447-50., .

Eid W, Steger M, El-Shemerly M, Ferretti LP, Pena-Diaz J, Konig C, Valtorta E, Sartori AA, and Ferrari S, 2010. DNA end resection by CtIP and exonuclease 1 prevents genomic instability. - *EMBO Rep.*2010 Dec;11(12):962-8.Epub 2010 Nov 5., .

El-Shemerly M, Hess D, Pyakurel AK, Moselhy S, and Ferrari S, 2007. ATR-dependent pathways control hEXO1 stability in response to stalled forks. - *Nucleic Acids Res.*2008 Feb;36(2):511-9.Epub 2007 Nov 29., .

El-Shemerly M, Janscak P, Hess D, Jiricny J, and Ferrari S, 2005. Degradation of human exonuclease 1b upon DNA synthesis inhibition. - *Cancer Res.*2005 May 1;65(9):3604-9., .

Fasching CL, Neumann AA, Muntoni A, Yeager TR, and Reddel RR, 2007. DNA damage induces alternative lengthening of telomeres (ALT) associated promyelocytic leukemia bodies that preferentially associate with linear telomeric DNA. - *Cancer Res.*2007 Aug 1;67(15):7072-7.Epub 2007 Jul 24., .

Feldmann E, Schmiemann V, Goedecke W, Reichenberger S, and Pfeiffer P, 2000. DNA double-strand break repair in cell-free extracts from Ku80-deficient cells: implications for Ku serving as an alignment factor in non-homologous DNA end joining. - *Nucleic Acids Res.*2000 Jul 1;28(13):2585-96., .

Felton KE, Gilchrist DM, and Andrew SE, 2007. Constitutive deficiency in DNA mismatch repair. - *Clin Genet.*2007 Jun;71(6):483-98., .

Flamand L, Komaroff AL, Arbuckle JH, Medveczky PG, and Ablashi DV, 2010. Review, part 1: Human herpesvirus-6-basic biology, diagnostic testing, and antiviral efficacy. - *J Med Virol.*2010 Sep;82(9):1560-8., .

Fusco C, Raymond A, and Zervos AS, 1998. Molecular cloning and characterization of a novel retinoblastoma-binding protein. - *Genomics.*1998 Aug 1;51(3):351-8., .

Gao J, Luo X, Tang K, Li X, and Li G, 2006. Epstein-Barr virus integrates frequently into chromosome 4q, 2q, 1q and 7q of Burkitt's lymphoma cell line (Raji). - *J Virol Methods.*2006 Sep;136(1-2):193-9.Epub 2006 Jun 23., .

Germann MW, Johnson CN, and Spring AM, 2010. Recognition of Damaged DNA: Structure and Dynamic Markers. - *Med Res Rev.*2012 May;32(3):659-83.doi: 10.1002/med.20226.Epub 2010 Nov 9., .

Gompels UA, Nicholas J, Lawrence G, Jones M, Thomson BJ, Martin ME, Efstathiou S, Craxton M, and Macaulay HA, 1995. The DNA sequence of human herpesvirus-6:

- structure, coding content, and genome evolution. - *Virology*.1995 May 10;209(1):29-51., .
- Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, and de Lange T, 1999. Mammalian telomeres end in a large duplex loop. - *Cell*.1999 May 14;97(4):503-14., .
- Griffiths PD, Ait-Khaled M, Bearcroft CP, Clark DA, Quaglia A, Davies SE, Burroughs AK, Rolles K, Kidd IM, Knight SN, Noibi SM, Cope AV, Phillips AN, and Emery VC, 1999. Human herpesviruses 6 and 7 as potential pathogens after liver transplant: prospective comparison with the effect of cytomegalovirus. - *J Med Virol*.1999 Dec;59(4):496-501., .
- Hackett JA, and Greider CW, 2003. End resection initiates genomic instability in the absence of telomerase. - *Mol Cell Biol*.2003 Dec;23(23):8450-61., .
- Hall CB, Caserta MT, Schnabel K, Shelley LM, Marino AS, Carnahan JA, Yoo C, Lofthus GK, and McDermott MP, 2008. Chromosomal integration of human herpesvirus 6 is the major mode of congenital human herpesvirus 6 infection. - *Pediatrics*.2008 Sep;122(3):513-20., .
- Harley CB, Futcher AB, and Greider CW, 1990. Telomeres shorten during ageing of human fibroblasts. - *Nature*.1990 May 31;345(6274):458-60., .
- Hayflick, L. & Moorhead, P., 1961. The serial cultivation of human diploid cell strains. - *Exp Cell Res*.1961 Dec;25:585-621., .
- Henson JD, Cao Y, Huschtscha LI, Chang AC, Au AY, Pickett HA, and Reddel RR, 2009. DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. - *Nat Biotechnol*.2009 Dec;27(12):1181-5.Epub ., .
- Henson JD, Neumann AA, Yeager TR, and Reddel RR, 2002. Alternative lengthening of telomeres in mammalian cells. - *Oncogene*.2002 Jan 21;21(4):598-610., .
- Heyer WD, Ehmsen KT, and Liu J, 2010. Regulation of homologous recombination in eukaryotes. - *Annu Rev Genet*.2010;44:113-39., .
- Holthausen JT, Wyman C, and Kanaar R, 2010. Regulation of DNA strand exchange in homologous recombination. - *DNA Repair (Amst)*.2010 Dec 10;9(12):1264-72., .
- Huertas P, Cortes-Ledesma F, Sartori AA, Aguilera A, and Jackson SP, 2008. CDK targets Sae2 to control DNA-end resection and homologous recombination. - *Nature*.2008 Oct 2;455(7213):689-92.Epub 2008 Aug 20., .
- Huertas P, and Jackson SP, 2009. Human CtIP mediates cell cycle control of DNA end resection and double strand break repair. - *J Biol Chem*.2009 Apr 3;284(14):9558-65.Epub 2009 Feb 7., .

Hurley EA, Agger S, McNeil JA, Lawrence JB, Calendar A, Lenoir G, and Thorley-Lawson DA, 1991. When Epstein-Barr virus persistently infects B-cell lines, it frequently integrates. - *J Virol.*1991 Mar;65(3):1245-54., .

Huser D, Gogol-Doring A, Lutter T, Weger S, Winter K, Hammer EM, Cathomen T, Reinert K, and Heilbronn R, 2010. Integration preferences of wildtype AAV-2 for consensus re-binding sites at numerous loci in the human genome. - *PLoS Pathog.*2010 Jul 8;6(7):e1000985., .

Iglesias N, and Lingner J, 2009. Related mechanisms for end processing at telomeres and DNA double-strand breaks. - *Mol Cell.*2009 Jul 31;35(2):137-8., .

Jagmohan-Changur S, Poikonen T, Vilkki S, Launonen V, Wikman F, Orntoft TF, Moller P, Vasen H, Tops C, Kolodner RD, Mecklin JP, Jarvinen H, Bevan S, Houlston RS, Aaltonen LA, 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 Res.*2003 Jan 1;63(1):154-8., .

Jeffreys AJ, Tamaki K, MacLeod A, Monckton DG, Neil DL, and Armour JA, 1994. Complex gene conversion events in germline mutation at human minisatellites. - *Nat Genet.*1994 Feb;6(2):136-45., .

Jeyapalan JN, Mendez-Bermudez A, Zaffaroni N, Dubrova YE, and Royle NJ, 2008. Evidence for alternative lengthening of telomeres in liposarcomas in the absence of ALT-associated PML bodies. - *Int J Cancer.*2008 Jun 1;122(11):2414-21., .

Jeyapalan JN, Varley H, Foxon JL, Pollock RE, Jeffreys AJ, Henson JD, Reddel RR, and Royle NJ, 2005. Activation of the ALT pathway for telomere maintenance can affect other sequences in the human genome. - *Hum Mol Genet.*2005 Jul 1;14(13):1785-94.Epub 2005 May 11., .

Jiang WQ, Zhong ZH, Henson JD, and Reddel RR, 2007. Identification of candidate alternative lengthening of telomeres genes by methionine restriction and RNA interference. - *Oncogene.*2007 Jul 12;26(32):4635-47.Epub 2007 Feb 5., .

Jiang WQ, Zhong ZH, Henson JD, Neumann AA, Chang AC, and Reddel RR, 2005. Suppression of alternative lengthening of telomeres by Sp100-mediated sequestration of the MRE11/RAD50/NBS1 complex. - *Mol Cell Biol.*2005 Apr;25(7):2708-21., .

Johnson FB, Marciniak RA, McVey M, Stewart SA, Hahn WC, and Guarente L, 2001. The *Saccharomyces cerevisiae* WRN homolog Sgs1p participates in telomere maintenance in cells lacking telomerase. - *EMBO J.*2001 Feb 15;20(4):905-13., .

Kaidi A, Weinert BT, Choudhary C, and Jackson SP, 2010. Human SIRT6 promotes DNA end

resection through CtIP deacetylation. - *Science*.2010 Sep 10;329(5997):1348-53., .

Kaufer BB, Jarosinski KW and Osterrieder N, 2011. Herpesvirus telomeric repeats facilitate genomic integration into host telomeres and mobilization of viral DNA during reactivation. - *J Exp Med*.2011 Mar 7., .

Kawabata A, Oyaizu H, Maeki T, Tang H, Yamanishi K, and Mori Y, 2011. Analysis of a neutralizing antibody for human herpesvirus 6B reveals a role for glycoprotein Q1 in viral entry. - *J Virol*.2011 Dec;85(24):12962-71.Epub 2011 Sep 28., .

Knudsen NO, Nielsen FC, Vinther L, Bertelsen R, Holten-Andersen S, Liberti SE, Hofstra R, Kooi K, and Rasmussen LJ, 2007. Nuclear localization of human DNA mismatch repair protein exonuclease 1 (hEXO1). - *Nucleic Acids Res*.2007;35(8):2609-19.Epub 2007 Apr 10., .

Kondo K, Sashihara J, Shimada K, Takemoto M, Amo K, Miyagawa H, and Yamanishi K, 2003. Recognition of a novel stage of betaherpesvirus latency in human herpesvirus 6. - *J Virol*.2003 Feb;77(3):2258-64., .

Kondo K, Shimada K, Sashihara J, Tanaka-Taya K, and Yamanishi K, 2002. Identification of human herpesvirus 6 latency-associated transcripts. - *J Virol*.2002 Apr;76(8):4145-51., .

Krejci, L., Altmannova, V., Spirek, M. and Zhao, X., 2012. Homologous recombination and its regulation. *Nucleic Acids Research*, .

Krueger GR, and Ablashi DV, 2003. Human herpesvirus-6: a short review of its biological behavior. - *Intervirology*.2003;46(5):257-69., .

Kyo S, Takakura M, Fujiwara T, and Inoue M, 2008. Understanding and exploiting hTERT promoter regulation for diagnosis and treatment of human cancers. - *Cancer Sci*.2008 Aug;99(8):1528-38., .

Laud, P.R., Multani, A.S., Bailey, S.M., Wu, L., Ma, J., Kingsley, C., Lebel, M., Pathak, S., DePinho, R.A. and Chang, S., 2005. Elevated telomere-telomere recombination in WRN-deficient, telomere dysfunctional cells promotes escape from senescence and engagement of the ALT pathway. *Genes Dev.*, **19**:2560-2570.

Lee Bi BI, Nguyen LH, Barsky D, Fernandes M, and Wilson DM 3rd, 2002. Molecular interactions of human Exo1 with DNA. - *Nucleic Acids Res*.2002 Feb 15;30(4):942-9., .

Lee BI, and Wilson DM 3rd, 1999. The RAD2 domain of human exonuclease 1 exhibits 5' to 3' exonuclease and flap structure-specific endonuclease activities. - *J Biol Chem*.1999 Dec 31;274(53):37763-9., .

Lee SO, Brown RA, and Razonable RR, 2011. Clinical significance of pretransplant chromosomally integrated human herpesvirus-6 in liver transplant recipients. - *Transplantation*.2011 Jul 27;92(2):224-9., .

Lee, K. & Lee, S.E., 2007. Saccharomyces cerevisiae Sae2- and Tel1-Dependent Single-Strand DNA Formation at DNA Break Promotes Microhomology-Mediated End Joining. *Genetics*, **176**2003-2014.

Lengsfeld BM, Rattray AJ, Bhaskara V, Ghirlando R, and Paull TT, 2007. Sae2 is an endonuclease that processes hairpin DNA cooperatively with the Mre11/Rad50/Xrs2 complex. - *Mol Cell*.2007 Nov 30;28(4):638-51., .

Leong HN, Tuke PW, Tedder RS, Khanom AB, Eglin RP, Atkinson CE, Ward KN, Griffiths PD, and Clark DA, 2007. The prevalence of chromosomally integrated human herpesvirus 6 genomes in the blood of UK blood donors. - *J Med Virol*.2007 Jan;79(1):45-51., .

Letsolo BT, Rowson J, and Baird DM, 2009. Fusion of short telomeres in human cells is characterized by extensive deletion and microhomology, and can result in complex rearrangements. - *Nucleic Acids Res*.2010 Apr;38(6):1841-52.Epub 2009 Dec 21., .

Li S, Ting NS, Zheng L, Chen PL, Ziv Y, Shiloh Y, Lee EY, and Lee WH, 2000. Functional link of BRCA1 and ataxia telangiectasia gene product in DNA damage response. - *Nature*.2000 Jul 13;406(6792):210-5., .

Liang L, Deng L, Nguyen SC, Zhao X, Maulion CD, Shao C, and Tischfield JA, 2008. Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks. - *Nucleic Acids Res*.2008 Jun;36(10):3297-310.Epub 2008 Apr 24., .

Lillard-Wetherell K, Machwe A, Languard GT, Combs KA, Behbehani GK, Schonberg SA, German J, Turchi JJ, Orren DK, and Groden J, 2004. Association and regulation of the BLM helicase by the telomere proteins TRF1 and TRF2. - *Hum Mol Genet*.2004 Sep 1;13(17):1919-32.Epub 2004 Jun 30., .

Liu F, and Lee WH, 2006. CtIP activates its own and cyclin D1 promoters via the E2F/RB pathway during G1/S progression. - *Mol Cell Biol*.2006 Apr;26(8):3124-34., .

Liu J, Doty T, Gibson B, and Heyer WD, 2011. Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. - *Nat Struct Mol Biol*.2010 Oct;17(10):1260-2.Epub 2010 Aug 22., .

Liu L, Saldanha SN, Pate MS, Andrews LG, and Tollefsbol TO, 2004. Epigenetic regulation of human telomerase reverse transcriptase promoter activity during cellular differentiation. -

Genes Chromosomes Cancer.2004 Sep;41(1):26-37., .

Llorente B, Smith CE, and Symington LS, 2008. Break-induced replication: what is it and what is it for? - *Cell Cycle*.2008 Apr 1;7(7):859-64.Epub 2008 Jan 14., .

Loayza D, and De Lange T, 2003. POT1 as a terminal transducer of TRF1 telomere length control. - *Nature*.2003 Jun 26;423(6943):1013-8.Epub 2003 May 25., .

Londono-Vallejo JA, Der-Sarkissian H, Cazes L, Bacchetti S, and Reddel RR, 2004. Alternative lengthening of telomeres is characterized by high rates of telomeric exchange. - *Cancer Res*.2004 Apr 1;64(7):2324-7., .

Lopatina NG, Poole JC, Saldanha SN, Hansen NJ, Key JS, Pita MA, Andrews LG, and Tollefsbol TO, 2003. Control mechanisms in the regulation of telomerase reverse transcriptase expression in differentiating human teratocarcinoma cells. - *Biochem Biophys Res Commun*.2003 Jul 4;306(3):650-9., .

Lord CJ, and Ashworth A, 2012. The DNA damage response and cancer therapy. - *Nature*.2012 Jan 18;481(7381):287-94.doi: 10.1038/nature10760., .

Lu CY, Tsai CH, Brill SJ, and Teng SC, 2010. Sumoylation of the BLM ortholog, Sgs1, promotes telomere-telomere recombination in budding yeast. - *Nucleic Acids Res*.2010 Jan;38(2):488-98.Epub 2009 Nov 11., .

Luo G, Santoro IM, McDaniel LD, Nishijima I, Mills M, Youssoufian H, Vogel H, Schultz RA, and Bradley A, 2000. Cancer predisposition caused by elevated mitotic recombination in Bloom mice. - *Nat Genet*.2000 Dec;26(4):424-9., .

Luppi M, Marasca R, Barozzi P, Ferrari S, Ceccherini-Nelli L, Batoni G, Merelli E, and Torelli G, 1993. Three cases of human herpesvirus-6 latent infection: integration of viral genome in peripheral blood mononuclear cell DNA. - *J Med Virol*.1993 May;40(1):44-52., .

Machwe A, Xiao L, Groden J, Matson SW, and Orren DK, 2005. RecQ family members combine strand pairing and unwinding activities to catalyze strand exchange. - *J Biol Chem*.2005 Jun 17;280(24):23397-407.Epub 2005 Apr 20., .

Maringele L, and Lydall D, 2004. EXO1 plays a role in generating type I and type II survivors in budding yeast. - *Genetics*.2004 Apr;166(4):1641-9., .

Marrero VA, and Symington LS, 2010. Extensive DNA end processing by exo1 and sgs1 inhibits break-induced replication. - *PLoS Genet*.2010 Jul 8;6(7):e1001007., .

Martens UM, Zijlmans JM, Poon SS, Dragowska W, Yui J, Chavez EA, Ward RK, and Lansdorp PM, 1998. Short telomeres on human chromosome 17p. - *Nat Genet*.1998

Jan;18(1):76-80., .

Mathieu N, Pirzio L, Freulet-Marriere MA, Desmaze C, and Sabatier L, 2004. Telomeres and chromosomal instability. - *Cell Mol Life Sci.*2004 Mar;*61(6):641-56., .*

McEachern MJ, and Haber JE, 2006. Break-induced replication and recombinational telomere elongation in yeast. - *Annu Rev Biochem.*2006;*75:111-35., .*

McVey, M. & Lee, S.E., 2008. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends in Genetics*, **24**:529-538.

Meyer WK, Arbeithuber B, Ober C, Ebner T, Tiemann-Boege I, Hudson RR, and Przeworski M, 2012. Evaluating the evidence for transmission distortion in human pedigrees. - *Genetics.*2012 May;*191(1):215-32.Epub 2012 Feb 29., .*

Michishita E, McCord RA, Berber E, Kioi M, Padilla-Nash H, Damian M, Cheung P, Kusumoto R, Kawahara TL, Barrett JC, Chang HY, Bohr VA, Ried T, Gozani O, and Chua KF, 2008. SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. - *Nature.*2008 Mar 27;*452(7186):492-6.Epub 2008 Mar 12., .*

Michou V, Liarmakopoulou S, Thomas D, Tsimaratou K, Makarounis K, Constantoulakis P, Angelopoulou R and Tsilivakos V, 2011. Herpes virus infected spermatozoa following density gradient centrifugation for IVF purposes. LID - 10.1111/j.1439-0272.2010.01121.x [doi]. - *Andrologia.*2011 Aug 16.doi: 10.1111/j.1439-0272.2010.01121.x., .

Mimitou EP, and Symington LS, 2008. Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. - *Nature.*2008 Oct 9;*455(7214):770-4.Epub 2008 Sep 21., .*

Mitchell JR, and Collins K, 2000. Human telomerase activation requires two independent interactions between telomerase RNA and telomerase reverse transcriptase. - *Mol Cell.*2000 Aug;*6(2):361-71., .*

Mitchell JR, Wood E, and Collins K, 1999. A telomerase component is defective in the human disease dyskeratosis congenita. - *Nature.*1999 Dec 2;*402(6761):551-5., .*

Monaghan P, 2010. Telomeres and life histories: the long and the short of it. - *Ann N Y Acad Sci.*2010 Sep;*1206:130-42.doi: 10.1111/j.1749-6632.2010.05705.x., .*

Mori Y, Dhepakson P, Shimamoto T, Ueda K, Gomi Y, Tani H, Matsuura Y, and Yamanishi K, 2000. Expression of human herpesvirus 6B rep within infected cells and binding of its gene product to the TATA-binding protein in vitro and in vivo. - *J Virol.*2000 Jul;*74(13):6096-104., .*

Morin I, Ngo HP, Greenall A, Zubko MK, Morrice N, and Lydall D, 2008. Checkpoint-

dependent phosphorylation of Exo1 modulates the DNA damage response. - *EMBO J.*2008 Sep 17;27(18):2400-10.Epub 2008 Aug 28., .

Morissette G, FAU - Flamand, L. and Flamand L, 2010. Herpesviruses and chromosomal integration. - *J Virol.*2010 Dec;84(23):12100-9.Epub 2010 Sep 15., .

Mostoslavsky R, Chua KF, Lombard DB, Pang WW, Fischer MR, Gellon L, Liu P, Mostoslavsky G, Franco S, Murphy MM, Mills KD, Patel P, Hsu JT, Hong AL, Ford E, Cheng HL, Kennedy C, Nunez N, Bronson R, Frendewey D, Auerbach W, Valenzuela D, Karow M, Hottiger MO, Hursting S, Barrett JC, Guarente L, Mulligan R, Demple B, Yancopoulos GD, and Alt FW, 2006. Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. - *Cell.*2006 Jan 27;124(2):315-29., .

Moyzis RK, Buckingham JM, Cram LS, Dani M, Deaven LL, Jones MD, Meyne J, Ratliff RL, and Wu JR, 1988. A highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. - *Proc Natl Acad Sci U S A.*1988 Sep;85(18):6622-6., .

Nabetani A, and Ishikawa F, 2009. Unusual telomeric DNAs in human telomerase-negative immortalized cells. - *Mol Cell Biol.*2009 Feb;29(3):703-13.Epub 2008 Nov 17., .

Nacheva EP, Ward KN, Brazma D, Virgili A, Howard J, Leong HN, and Clark DA, 2008. Human herpesvirus 6 integrates within telomeric regions as evidenced by five different chromosomal sites. - *J Med Virol.*2008 Nov;80(11):1952-8., .

Nimonkar AV, Genschel J, Kinoshita E, Polaczek P, Campbell JL, Wyman C, Modrich P, and Kowalczykowski SC, 2011. BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. - *Genes Dev.*2011 Feb 15;25(4):350-62., .

Nimonkar AV, Ozsoy AZ, Genschel J, Modrich P, and Kowalczykowski SC, 2008. Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair. - *Proc Natl Acad Sci U S A.*2008 Nov 4;105(44):16906-11.Epub 2008 Oct 29., .

Nittis T, Guittat L, and Stewart SA, 2008. Alternative lengthening of telomeres (ALT) and chromatin: is there a connection? - *Biochimie.*2008 Jan;90(1):5-12.Epub 2007 Sep 2., .

Ochman H, Gerber AS, and Hartl DL, 1988. Genetic applications of an inverse polymerase chain reaction. - *Genetics.*1988 Nov;120(3):621-3., .

Ogata M, 2009. Human herpesvirus 6 in hematological malignancies. - *J Clin Exp Hematop.*2009 Nov;49(2):57-67., .

Okuno T, Takahashi K, Balachandra K, Shiraki K, Yamanishi K, Takahashi M, and Baba K, 1989. Seroepidemiology of human herpesvirus 6 infection in normal children and adults. -

*J Clin Microbiol.*1989 Apr;27(4):651-3., .

Olovnikov AM, 1973. A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. - *J Theor Biol.*1973 Sep 14;41(1):181-90., .

Opresko PL, Mason PA, Podell ER, Lei M, Hickson ID, Cech TR, and Bohr VA, 2005. POT1 stimulates RecQ helicases WRN and BLM to unwind telomeric DNA substrates. - *J Biol Chem.*2005 Sep 16;280(37):32069-80.Epub 2005 Jul 18., .

Opresko PL, von Kobbe C, Laine JP, Harrigan J, Hickson ID, and Bohr VA, 2002. Telomere-binding protein TRF2 binds to and stimulates the Werner and Bloom syndrome helicases. - *J Biol Chem.*2002 Oct 25;277(43):41110-9.Epub 2002 Aug 13., .

Palm W, and de Lange T, 2008. How shelterin protects Mammalian telomeres. - *Annu Rev Genet.*2008;42:301-34., .

Pardo B, Gomez-Gonzalez B, and Aguilera A, 2009. DNA repair in mammalian cells: DNA double-strand break repair: how to fix a broken relationship. - *Cell Mol Life Sci.*2009 Mar;66(6):1039-56., .

Pellett PE, Ablashi DV, Ambros PF, Agut H, Caserta MT, Descamps V, Flamand L, Gautheret-Dejean A, Hall CB, Kamble RT, Kuehl U, Lassner D, Lautenschlager I, Loomis KS, Luppi M, Lusso P, Medveczky PG, Montoya JG, Mori Y, Ogata M, Pritchett JC, Rogez S, Seto E, Ward KN, Yoshikawa T and Razonable RR, 2011. Chromosomally integrated human herpesvirus 6: questions and answers. LID - 10.1002/rmv.715 [doi]. - *Rev Med Virol.*2011 Nov 4.doi: 10.1002/rmv.715., .

Pickett HA, Cesare AJ, Johnston RL, Neumann AA, and Reddel RR, 2009. Control of telomere length by a trimming mechanism that involves generation of t-circles. - *EMBO J.*2009 Apr 8;28(7):799-809.Epub 2009 Feb 12., .

Pickett HA, Henson JD, Au AY, Neumann AA, and Reddel RR, 2011. Normal mammalian cells negatively regulate telomere length by telomere trimming. - *Hum Mol Genet.*2011 Dec 1;20(23):4684-92.Epub 2011 Sep 8., .

Qvist P, Huertas P, Jimeno S, Nyegaard M, Hassan MJ, Jackson SP, and Borglum AD, 2011. CtIP Mutations Cause Seckel and Jawad Syndromes. - *PLoS Genet.*2011 Oct;7(10):e1002310.Epub 2011 Oct 6., .

Rageul J, Fremin C, Ezan F, Baffet G, and Langouet S, 2011. The knock-down of ERCC1 but not of XPF causes multinucleation. - *DNA Repair (Amst).*2011 Sep 5;10(9):978-90.Epub 2011 Aug 12., .

Razzaque A, 1990. Oncogenic potential of human herpesvirus-6 DNA. - *Oncogene*.1990 Sep;5(9):1365-70., .

Reddel RR, 2003. Alternative lengthening of telomeres, telomerase, and cancer. - *Cancer Lett*.2003 May 15;194(2):155-62., .

Reddel RR, Bryan TM, Colgin LM, Perrem KT, and Yeager TR, 2001. Alternative lengthening of telomeres in human cells. - *Radiat Res*.2001 Jan;155(1 Pt 2):194-200.

Chen, Q., Ijpm, A. and Greider, C.W., 2001. Two survivor pathways that allow growth in the absence of telomerase are generated by distinct telomere recombination events. *Mol.Cell.Biol.*, **21**1819-1827.

Reddel, R.R., Bryan, T.M. and Murnane, J.P., 1997. Immortalized cells with no detectable telomerase activity. A review. *Biochemistry (Mosc)*, **62**1254-1262.

Renaud S, Loukinov D, Abdullaev Z, Guilleret I, Bosman FT, Lobanenkova V, and Benhattar J, 2007. Dual role of DNA methylation inside and outside of CTCF-binding regions in the transcriptional regulation of the telomerase hTERT gene. - *Nucleic Acids Res*.2007;35(4):1245-56.Epub 2007 Jan 31., .

Riethman H, 2008. Human telomere structure and biology. - *Annu Rev Genomics Hum Genet*.2008;9:1-19., .

Rossi ML, Ghosh AK, and Bohr VA, 2010. Roles of Werner syndrome protein in protection of genome integrity. - *DNA Repair (Amst)*.2010 Mar 2;9(3):331-44.Epub 2010 Jan 13., .

Rotola A, Ravaioli T, Gonelli A, Dewhurst S, Cassai E, and Di Luca D, 1998. U94 of human herpesvirus 6 is expressed in latently infected peripheral blood mononuclear cells and blocks viral gene expression in transformed lymphocytes in culture. - *Proc Natl Acad Sci U S A*.1998 Nov 10;95(23):13911-6., .

Ryan MC, Zeeberg BR, Caplen NJ, Cleland JA, Kahn AB, Liu H, and Weinstein JN, 2008. SpliceCenter: a suite of web-based bioinformatic applications for evaluating the impact of alternative splicing on RT-PCR, RNAi, microarray, and peptide-based studies. - *BMC Bioinformatics*.2008 Jul 18;9:313., .

Saharia A & Stewart SA, 2009. FEN1 contributes to telomere stability in ALT-positive tumor cells. - *Oncogene*.2009 Jan 12., .

Salahuddin SZ, Ablashi DV, Markham PD, Josephs SF, Sturzenegger S, Kaplan M, Halligan G, Biberfeld P, Wong-Staal F, and Kramarsky B, 1986. Isolation of a new virus, HBLV, in patients with lymphoproliferative disorders. - *Science*.1986 Oct 31;234(4776):596-601., .

Sartori AA, Lukas C, Coates J, Mistrik M, Fu S, Bartek J, Baer R, Lukas J, and Jackson SP, 2007. Human CtIP promotes DNA end resection. - *Nature*.2007 Nov 22;450(7169):509-14.Epub 2007 Oct 28., .

Savage SA, and Bertuch AA, 2010. The genetics and clinical manifestations of telomere biology disorders. - *Genet Med*.2010 Dec;12(12):753-64., .

Schaeper U, Subramanian T, Lim L, Boyd JM, and Chinnadurai G, 1998. Interaction between a cellular protein that binds to the C-terminal region of adenovirus E1A (CtBP) and a novel cellular protein is disrupted by E1A through a conserved PLDLS motif. - *J Biol Chem*.1998 Apr 10;273(15):8549-52., .

Schaetzlein S, Kodandaramireddy NR, Ju Z, Lechel A, Stepczynska A, Lilli DR, Clark AB, Rudolph C, Kuhnel F, Wei K, Schlegelberger B, Schirmacher P, Kunkel TA, Greenberg RA, Edelmann W, and Rudolph KL, 2007. Exonuclease-1 deletion impairs DNA damage signaling and prolongs lifespan of telomere-dysfunctional mice. - *Cell*.2007 Sep 7;130(5):863-77., .

Schmutte C, Sadoff MM, Shim KS, Acharya S, and Fishel R, 2001. The interaction of DNA mismatch repair proteins with human exonuclease I. - *J Biol Chem*.2001 Aug 31;276(35):33011-8.Epub 2001 Jun 26., .

Sharma S, Otterlei M, Sommers JA, Driscoll HC, Dianov GL, Kao HI, Bambara RA, and Brosh RM Jr, 2004. WRN helicase and FEN-1 form a complex upon replication arrest and together process branchmigrating DNA structures associated with the replication fork. - *Mol Biol Cell*.2004 Feb;15(2):734-50.Epub 2003 Dec 2., .

Sharma S, Sommers JA, Driscoll HC, Uzdilla L, Wilson TM, and Brosh RM Jr, 2003. The exonucleolytic and endonucleolytic cleavage activities of human exonuclease 1 are stimulated by an interaction with the carboxyl-terminal region of the Werner syndrome protein. - *J Biol Chem*.2003 Jun 27;278(26):23487-96.Epub 2003 Apr 18., .

Shay JW, and Bacchetti S, 1997. A survey of telomerase activity in human cancer. - *Eur J Cancer*.1997 Apr;33(5):787-91., .

Shay JW, and Roninson IB, 2004. Hallmarks of senescence in carcinogenesis and cancer therapy. - *Oncogene*.2004 Apr 12;23(16):2919-33., .

Shiloh Y, 2003. ATM and related protein kinases: safeguarding genome integrity. - *Nat Rev Cancer*.2003 Mar;3(3):155-68., .

Smogorzewska A, and de Lange T, 2004. Regulation of telomerase by telomeric proteins. - *Annu Rev Biochem*.2004;73:177-208., .

Sugiyama T, Kantake N, Wu Y, and Kowalczykowski SC, 2006. Rad52-mediated DNA annealing after Rad51-mediated DNA strand exchange promotes second ssDNA capture. - *EMBO J.*2006 Nov 29;25(23):5539-48.Epub 2006 Nov 9., .

Sun X, Zheng L, and Shen B, 2002. Functional alterations of human exonuclease 1 mutants identified in atypical hereditary nonpolyposis colorectal cancer syndrome. - *Cancer Res.*2002 Nov 1;62(21):6026-30., .

Suwaki N, Klare K, and Tarsounas M, 2011. RAD51 paralogs: roles in DNA damage signalling, recombinational repair and tumorigenesis. - *Semin Cell Dev Biol.*2011 Oct;22(8):898-905.Epub 2011 Jul 28., .

Symington LS, and Gautier J, 2011. Double-strand break end resection and repair pathway choice. - *Annu Rev Genet.*2011;45:247-71.Epub 2011 Sep 12., .

Takai H, Smogorzewska A, and de Lange T, 2003. DNA damage foci at dysfunctional telomeres. - *Curr Biol.*2003 Sep 2;13(17):1549-56., .

Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, Utsumi H, Yamaguchi-Iwai Y, Shinohara A, and Takeda S, 1998. Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. - *EMBO J.*1998 Sep 15;17(18):5497-508., .

Tanaka-Taya K, Sashihara J, Kurahashi H, Amo K, Miyagawa H, Kondo K, Okada S, and Yamanishi K, 2004. Human herpesvirus 6 (HHV-6) is transmitted from parent to child in an integrated form and characterization of cases with chromosomally integrated HHV-6 DNA. - *J Med Virol.*2004 Jul;73(3):465-73., .

Temime-Smaali N, Guittat L, Sidibe A, Shin-ya K, Trentesaux C, and Riou JF, 2009. The G-quadruplex ligand telomestatin impairs binding of topoisomerase IIIalpha to G-quadruplex-forming oligonucleotides and uncaps telomeres in ALT cells. - *PLoS One.*2009 Sep 9;4(9):e6919., .

Temime-Smaali N, Guittat L, Wenner T, Bayart E, Douarre C, Gomez D, Giraud-Panis MJ, Londono-Vallejo A, Gilson E, Amor-Gueret M, and Riou JF, 2008. Topoisomerase IIIalpha is required for normal proliferation and telomere stability in alternative lengthening of telomeres. - *EMBO J.*2008 May 21;27(10):1513-24.Epub 2008 Apr 17., .

Tennen RI, and Chua KF, 2011. Chromatin regulation and genome maintenance by mammalian SIRT6. - *Trends Biochem Sci.*2011 Jan;36(1):39-46.Epub 2010 Aug 21., .

Thomson BJ, Weindler FW, Gray D, Schwaab V, and Heilbronn R, 1994. Human herpesvirus 6 (HHV-6) is a helper virus for adeno-associated virus type 2 (AAV-2) and the

AAV-2 rep gene homologue in HHV-6 can mediate AAV-2 DNA replication and regulate gene expression. - *Virology*.1994 Oct;204(1):304-11., .

Tishkoff DX, Amin NS, Viars CS, Arden KC, and Kolodner RD, 1998. Identification of a human gene encoding a homologue of *Saccharomyces cerevisiae* EXO1, an exonuclease implicated in mismatch repair and recombination. - *Cancer Res*.1998 Nov 15;58(22):5027-31., .

Tomaska L, Nosek J, Kramara J, and Griffith JD, 2009. Telomeric circles: universal players in telomere maintenance? - *Nat Struct Mol Biol*.2009 Oct;16(10):1010-5.Epub 2009 Oct 6., .

Tong L, and Stow ND, 2010. Analysis of herpes simplex virus type 1 DNA packaging signal mutations in the context of the viral genome. - *J Virol*.2010 Jan;84(1):321-9.Epub ., .

Torelli G, Barozzi P, Marasca R, Cocconcelli P, Merelli E, Ceccherini-Nelli L, Ferrari S, and Luppi M, 1995. Targeted integration of human herpesvirus 6 in the p arm of chromosome 17 of human peripheral blood mononuclear cells in vivo. - *J Med Virol*.1995 Jul;46(3):178-88., .

Tran PT, Erdeniz N, Symington LS, and Liskay RM, 2004. EXO1-A multi-tasking eukaryotic nuclease. - *DNA Repair (Amst)*.2004 Dec 2;3(12):1549-59., .

Tran PT, Fey JP, Erdeniz N, Gellon L, Boiteux S, and Liskay RM, 2007. A mutation in EXO1 defines separable roles in DNA mismatch repair and post-replication repair. - *DNA Repair (Amst)*.2007 Nov;6(11):1572-83.Epub 2007 Jun 29., .

Uniprot, . *Exonuclease 1 - Homo sapiens (Human)*. [online]. Available at: <http://www.uniprot.org/uniprot/Q9UQ84> [accessed 21/11/2008 2008].

Vallur AC, and Maizels N, 2010. Distinct activities of exonuclease 1 and flap endonuclease 1 at telomeric g4 DNA. - *PLoS One*.2010 Jan 26;5(1):e8908., .

van Geel M, Dickson MC, Beck AF, Bolland DJ, Frants RR, van der Maarel SM, de Jong PJ, and Hewitt JE, 2002. Genomic analysis of human chromosome 10q and 4q telomeres suggests a common origin. - *Genomics*.2002 Feb;79(2):210-7., .

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, and Speleman F, 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. - *Genome Biol*.2002 Jun 18;3(7):RESEARCH0034.Epub 2002 Jun 18., .

Varley H, Pickett HA, Foxon JL, Reddel RR, and Royle NJ, 2002. Molecular characterization of inter-telomere and intra-telomere mutations in human ALT cells. - *Nat Genet*.2002

Mar;30(3):301-5.Epub 2002 Feb 4., .

Vasa-Nicotera M, Brouillette S, Mangino M, Thompson JR, Braund P, Clemitson JR, Mason A, Bodycote CL, Raleigh SM, Louis E, and Samani NJ, 2005. Mapping of a major locus that determines telomere length in humans. - *Am J Hum Genet.*2005 Jan;*76(1):147-51.Epub 2004 Nov 1., .*

von Kobbe C, Karmakar P, Dawut L, Opresko P, Zeng X, Brosh RM Jr, Hickson ID, and Bohr VA, 2002. Colocalization, physical, and functional interaction between Werner and Bloom syndrome proteins. - *J Biol Chem.*2002 Jun *14;277(24):22035-44.Epub 2002 Mar 27., .*

Ward KN, Leong HN, Nacheva EP, Howard J, Atkinson CE, Davies NW, Griffiths PD, and Clark DA, 2006. Human herpesvirus 6 chromosomal integration in immunocompetent patients results in high levels of viral DNA in blood, sera, and hair follicles. - *J Clin Microbiol.*2006 Apr;*44(4):1571-4., .*

Wei W, and Sedivy JM, 1999. Differentiation between senescence (M1) and crisis (M2) in human fibroblast cultures. - *Exp Cell Res.*1999 Dec *15;253(2):519-22., .*

Wilson DM 3rd, Carney JP, Coleman MA, Adamson AW, Christensen M, and Lamerdin JE, 1998. Hex1: a new human Rad2 nuclease family member with homology to yeast exonuclease 1. - *Nucleic Acids Res.*1998 Aug *15;26(16):3762-8., .*

Wilson DM 3rd, and Thompson LH, 2007. Molecular mechanisms of sister-chromatid exchange. - *Mutat Res.*2007 Mar *1;616(1-2):11-23.Epub 2006 Dec 6., .*

Wong AK, Ormonde PA, Pero R, Chen Y, Lian L, Salada G, Berry S, Lawrence Q, Dayananth P, Ha P, Tavtigian SV, Teng DH, and Bartel PL, 1988. Characterization of a carboxy-terminal BRCA1 interacting protein. - *Oncogene.*1988 Nov *5;17(18):2279-85., .*

Wu L, and Hickson ID, 2006. DNA helicases required for homologous recombination and repair of damaged replication forks. - *Annu Rev Genet.*2006;*40:279-306., .*

Wu Y, Berends MJ, Post JG, Mensink RG, Verlind E, Van Der Sluis T, Kempinga C, Sijmons RH, van der Zee AG, Hollema H, Kleibeuker JH, Buys CH, and Hofstra RM, 2001. Germline mutations of EXO1 gene in patients with hereditary nonpolyposis colorectal cancer (HNPCC) and atypical HNPCC forms. - *Gastroenterology.*2001 Jun;*120(7):1580-7., .*

Wyllie FS, Jones CJ, Skinner JW, Haughton MF, Wallis C, Wynford-Thomas D, Faragher RG, and Kipling D, 2000. Telomerase prevents the accelerated cell ageing of Werner syndrome fibroblasts. - *Nat Genet.*2000 Jan;*24(1):16-7., .*

Wyman C, and Kanaar R, 2004. Homologous recombination: down to the wire. - *Curr*

*Biol.*2004 Aug 10;14(15):R629-31., .

Yamamoto H, Hanafusa H, Ouchida M, Yano M, Suzuki H, Murakami M, Aoe M, Shimizu N, Nakachi K, and Shimizu K, 2005. Single nucleotide polymorphisms in the EXO1 gene and risk of colorectal cancer in a Japanese population. - *Carcinogenesis*.2005 Feb;26(2):411-6.Epub 2004 Nov 18., .

Yamanishi K, Okuno T, Shiraki K, Takahashi M, Kondo T, Asano Y, and Kurata T, 1988. Identification of human herpesvirus-6 as a causal agent for exanthem subitum. - *Lancet*.1988 May 14;1(8594):1065-7., .

Yeager TR, Neumann AA, Englezou A, Huschtscha LI, Noble JR, and Reddel RR, 1999. Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. - *Cancer Res*.1999 Sep 1;59(17):4175-9., .

Yehezkel S, Segev Y, Viegas-Pequignot E, Skorecki K, and Selig S, 2008. Hypomethylation of subtelomeric regions in ICF syndrome is associated with abnormally short telomeres and enhanced transcription from telomeric regions. - *Hum Mol Genet*.2008 Sep 15;17(18):2776-89.Epub 2008 Jun 16., .

Zhang A, Zheng C, Hou M, Lindvall C, Li KJ, Erlandsson F, Bjorkholm M, Gruber A, Blennow E, and Xu D, 2003. Deletion of the telomerase reverse transcriptase gene and haploinsufficiency of telomere maintenance in Cri du chat syndrome. - *Am J Hum Genet*.2003 Apr;72(4):940-8.Epub 2003 Mar 10., .

Zhao Y, Hoshiyama H, Shay JW, and Wright WE, 2007. Quantitative telomeric overhang determination using a double-strand specific nuclease. - *Nucleic Acids Res*.2008 Feb;36(3):e14.Epub 2007 Dec 10., .

Zhong Q, Chen CF, Chen PL, and Lee WH, 2002. BRCA1 facilitates microhomology-mediated end joining of DNA double strand breaks. - *J Biol Chem*.2002 Aug 9;277(32):28641-7.Epub 2002 May 30., .

Zhong ZH, Jiang WQ, Cesare AJ, Neumann AA, Wadhwa R, and Reddel RR, 2007. Disruption of telomere maintenance by depletion of the MRE11/RAD50/NBS1 complex in cells that use alternative lengthening of telomeres. - *J Biol Chem*.2007 Oct 5;282(40):29314-22.Epub 2007 Aug 9., .

Zhu J, Zhao Y, and Wang S, 2010. Chromatin and epigenetic regulation of the telomerase reverse transcriptase gene. - *Protein Cell*.2010 Jan;1(1):22-32.Epub 2010 Feb 7., .

Zhu XD, Kuster B, Mann M, Petrini JH, and de Lange T, 2000. Cell-cycle-regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres. - *Nat Genet*.2000

Jul;25(3):347-52., .

Zhu Z, Chung WH, Shim EY, Lee SE, and Ira G, 2008. Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. - *Cell.2008 Sep 19;134(6):981-94., .*

Zvereva MI, Shcherbakova DM, and Dontsova OA, 2010. Telomerase: structure, functions, and activity regulation. - *Biochemistry (Mosc).2010 Dec;75(13):1563-83., .*