INVESTIGATING THE ROLE OF ATRX IN TELOMERIC STABILITY AND ESCAPE FROM CRISIS

A thesis submitted for the degree of Doctor of Philosophy (PhD) by

Helene E B Geiller

February 2020

Division of Cancer and Genetics

School of Medicine

Cardiff University

SUMMARY

Cancer cells gain replicative immortality through the upregulation of telomere maintenance mechanisms, and 10-15% of tumours activate the alternative lengthening of telomeres (ALT). The *ATRX* gene has been strongly linked to ALT and is often mutated in cancers that upregulate this pathway. The aim of this project was to establish the effect the loss of the *ATRX* gene and telomere dysfunction have on the initiation of the ALT phenotype.

The *ATRX* gene was knocked out in two human cell types: epithelial cancer cells and primary fibroblasts. Upon loss of ATRX and the induction of telomere stress, cells successfully escaped a telomere-driven crisis in both models. In the fibroblast cultures, clones exclusively immortalised by the upregulation and maintenance of ALT. The loss of ATRX led to telomere length heterogeneity pre-crisis, providing evidence that it plays a role in the maintenance of telomere chromatin. In contrast, in epithelial cells, upon the loss of telomerase and the induction of a telomere-driven crisis, ALT-like elongation was observed at 5 chromosome ends in 4.7% of escapees for which the insertion length appeared to be chromosome and allele specific. A further 18% of clones that died during crisis, displayed initiation of an ALT-like process, providing evidence that epithelial cells can transiently switch to ALT, but ultimately require telomerase activity for long-term survival. Finally, telomeres from ALT clones were sequenced and showed an increase in non-canonical variant repeats suggesting the use of telomeric DNA as template for elongation.

In conclusion, the loss of ATRX combined with telomere stress during crisis, is sufficient to initiate the ALT mechanism, and this study provides an insight into the transition from normal cells to malignancy upon the activation of the ALT mechanism, thus providing a tractable model to gain a better understanding of the ALT pathway.

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Firstly, I would like to thank Duncan for allowing me to take on this project and for guiding and encouraging me throughout my time in his lab. I have gained invaluable experience and enjoyed our supervisory meetings as they were a time to reflect on the project and reaffirmed my love for research. Thank you to Cardiff University and Cancer Research UK for funding my work.

I would also like to thank the entire STELA group: Alys, Isabel, Kate, Kez, Kev, Greg and especially Julia and Rhi, who not only trained me in the lab to be meticulous and adopt good habits for future work, but also undertook the preliminary work which enabled me to undertake this project. They also answered all my questions, and there were a lot, and were very supportive for which I will always be grateful. Kez's patience throughout the bioinformatics analysis will also be remembered as I had no prior experience. Finally, I am grateful that Alys and I started at the same time, as I had somebody to go through the motions with as well as share the moments of doubts, especially when something was not working!

A big thank you to the second floor of the Cancer and Genetics building especially Jo, Pete, Tom, Trish and Stefan with whom great times were spent during tea and lunch breaks and nights out. The quantity of cake available in the tower room will definitely be missed alongside everybody's excellent baking skills! I will miss working with you all.

Merci papa pour ton soutien et tes encouragements dans tout ce que j'entreprends et surtout pendant toutes mes études, 8 ans ça fait long. Merci d'être la quand j'ai un doute et pour tout ce que tu fais pour moi. Thanks to my supportive sister Marie, who was always at the end of a phone call when I needed it and who, together with Guillaume, showed great interest in my research, which I appreciate greatly! Thanks to my big brother, Yannick, and Marie, who sent me supportive messages and welcomed me into their home whenever I needed a break from Cardiff. Thanks to my wonderful niece and nephew, Juliette and Thomas, who forced me to disconnect from work when I went back to France and who kept me busy and entertained.

And finally, but not least, I would like to thank Tom who was incredibly supportive of my decision to do a PhD and who moved to Cardiff for us to start a life together. You were incredibly patient and understanding of my perpetual high levels of stress and my inability to switch off at times, as well as always being up for doing things at weekends and exploring lovely Wales.

I would like to dedicate this work to my mum who lost her battle with cancer in 2006, and who is my motivation and the reason for my determination and ambition to pursue work in cancer research.

ABBREVIATIONS

ADD	ATRX-DNMT3-DNMTL domain
ALT	Alternative lengthening of telomeres
APBs	ALT-associated PML bodies
APS	Ammonium persulfate
AT	Ataxia telangiectasia
ATM	Ataxia telangiectasia mutated protein
ATP	Adenosine triphosphate
ATR	Ataxia telangiectasia and Rad3-related protein
ATRX	Alpha-thalassemia mental retardation syndrome X-linked
AU	Arbitrary unit
BIR	Break-induced replication
BLAST	Basic local alignment search tool
BLM	Bloom protein
bp	Base pair
BRCA 1 and 2	Breast cancer gene 1 and 2
BSA	Bovine serum albumin
CCS	Circular consensus read
CENP-A	Centromeric protein A
cGAMP	Cyclic GMP-AMP
cGAS	Cyclic GMP-AMP synthase
CHAPS	3-cholamidopropyl dimethylammonio 1-propanesulfonate
Ci/mmol	Curies per millimole
CLES	Common language effect size
cm	Centimetres
CNS	Central nervous system
CNV	Copy number variant
CO_2	Carbon dioxide
CRISPR/Cas9	Clustered regularly interspaced short palindromic repeats/Caspase 9
Csf2ra	Colony stimulating factor 2 receptor alpha subunit
Ctrls	Controls
d-loop	Displacement loop
DAPI	4’,6-diamidino-2-phenylindole
DAXX	Death domain associated protein
DDR	DNA-damage response
dH_2O and ddH_2O	Distilled and double distilled water
DHB	DAXX helical bundle
Dhrsx	Dehydrogenase/reductase X-linked protein
DMEM	Dulbecco’s modified Eagle’s medium
DMSO	Dimethyl sulfoxide
DN-hTERT	Dominant negative hTERT

DNA-PK	DNA-dependent protein kinase
DNA	Deoxyribose nucleic acid
DNMT3	DNA methyltransferase 3
dNTP	Deoxyribonucleotide triphosphate
DSB	Double stranded break
DSBR	Double stranded break repair
dsDNA	Double stranded DNA
DTT	Dithiothreitol
ECL	Enhanced chemiluminescent
ECTRs	Extra chromosomal telomere repeats
EDTA	Ethylenediamine tetraacetic acid
EGTA	Ethyleneglycol tetraacetic
ERCC1	Excision repair cross-complementation group 1 protein
est1	Ever shorter telomeres protein 1
EtOH	Ethyl alcohol
EXO1	Exonuclease 1 protein
EZH2	Enhancer of zeste homolog 2
FACS	Fluorescence-activated cell sorting
FANCM	Fanconi anaemia complementation group M
FCS	Foetal calf serum
FISH	Fluorescent in situ hybridisation
G1 phase	Growth phase 1
G2 phase	Growth phase 2
G4	Guanine quadruplex
GFP	Green fluorescent protein
H2AX	H2A histone family member X
H3K4me0	Histone 3 unmethylated at lysine 4
H3K9me3	Histone 3 tri-methylated at lysine 9
H3K27me3	Histone 3 tri-methylated at lysine 27
HCl	Hydrochloric acid
HDR	Homology-directed repair
HER-2	Human epidermal growth factor receptor 2
HIRA	Histone cell cycle regulator
HMM	Hidden Markov model
HR	Homologous recombination
HP1	Heterochromatin protein 1
HPV	Human Papilloma virus
HRP	Horseradish peroxidase
hTERC	Telomerase RNA component
hTERT	Human telomere reverse transcriptase
IRF3	Interferon regulatory factor 3
ITS	Interstitial telomere sequences
kb	Kilobase
KCl	Potassium chloride

kDa	Kilo Dalton
KO	Knock-out
LFS	Li-Fraumeni syndrome
M	Molarity
M phase	Mitosis phase
MeCP2	Methyl CpG binding protein 2
MFH	Malignant fibrous histiocytoma
mg	Milligrams
$MgCl_2$	Magnesium chloride
mH2A	Macro H2A histone
ml	Millilitres
mM	Millimolar
MP	Mixed population
MRE11	Meiotic recombination 11 protein
MRN	MRE11, RAD50, NBS1 protein complex
mRNA	Messenger ribonucleic acid
MS32	Minisatellite 32
MSC	Mesenchymal stem cells
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NBS1	Nijmegen breakage syndrome 1 protein
ng	Nanograms
NGS	Next-generation sequencing
$(NH_4)SO_4$	Ammonium sulphate
NHEJ	Non-homologous end joining
Nlgn4	Neuroligin 4 protein
nM	Nanomolar
nm	Nanometres
PanNET	Primary pancreatic neuro-endocrine tumour
PBS	Phosphate-buffered saline
PCH	Pericentric heterochromatin
PCR	Polymerase chain reaction
PD	Population doubling
pg	Picogram
pH	Potential hydrogen
pM	Picomolar
PML	Pro-myelocytic leukaemia protein
PML NB	Pro-myelocytic protein nuclear body
PMSF	Phenylmethylsulfonyl fluoride
POT1	Protection of telomeres 1
PRC2	Polycomb repressive complex 2
PVDF	Polyvinylidene fluoride
PWO	Pyrococcus woesei enzyme
QC	Quality control

rAAV	Recombinant adeno-associated virus
RAP1	Repressor/Activator protein 1
Rb	Retinoblastoma
RCA	Rolling-circle amplification
RISC	RNA-inducing silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNase H1	Ribonuclease H1 protein
RPA	Replication protein A
rpm	Rotations per minute
RPMI	Roswell Park Memorial Institute
ROS	Reactive oxygen specie
S phase	Synthesis phase
SASP	Senescent-associated secretory phenotype
SCE	Sister chromatid exchanges
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SDSA	Synthesis-dependent strand annealing
shRNA	Short hairpin RNA
SMARCAL1	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A-like protein 1
SMRT	Single-molecule real-time
SNP	Single-nucleotide polymorphism
SSA	Single-strand annealing
SSC	Saline-sodium citrate
ssDNA	Single-strand DNA
STELA	Single telomere length analysis
STING	Stimulator of interferon genes
STS	Soft tissue sarcoma
SV	Structural variant
SWI2/SNF2	Switch 2/Sucrose non fermentable 2
t-loop	Telomeric loop
TAE buffer	Tris-acetate-EDTA buffer
TBE buffer	Tris-borate-EDTA buffer
TE buffer	Tris-EDTA buffer
TEMED	Tetramethylethylenediamine
TEN buffer	Tris-EDTA-NaCl buffer
TERRA	Telomeric repeat-containing RNA
TIN2	TRF1 and TRF2 interacting nuclear protein 2
TMM	Telomere maintenance mechanism
TP53	Tumour protein 53
TPG	Total product generated
TPP1	Tripeptidyl peptidase 1
TRAP	Telomere repeat amplification protocol

TRF	Terminal restriction fragment
TRF1 and 2	Telomeric repeat factor 1 and 2
TRIS	Trisaminomethane
trt-1	Telomerase reverse transcriptase 1
TVR	Telomere variant repeat
µg	Micrograms
µl	Microlitres
µM	Micromolar
UV	Ultra-violet
V	Volts
v/v	Volume per volume
VAF	Variant allele frequency
WGS	Whole genome sequencing
WRN	Werner protein
WT	Wild-type
ZMW	Zero-mode waveguides

Chapter 1

Introduction

1.1 Telomere Biology

1.1.1 Telomere structure

The ends of chromosomes were first described in 1930 by Hermann Muller who gave them their distinctive name of telomeres originating from the Greek; '*telo*' meaning end; and '*mere*' meaning part. He uncovered the property that they were resistant to ionising radiation and must therefore play a role in the protection of the natural ends of chromosomes (Muller 1938). It was not until 1978, that Elizabeth Blackburn discovered the repetitive nature of the telomeric sequence, present at the end of the ribosomal DNA in *Tetrahymena*, and that this was conserved in eukaryotes (Blackburn and Gall 1978). We now know that the telomere sequence is G-rich and varies depending on the specie: TTAGGG in mammalians; TTGGGG in *Tetrahymena*; TTTAGGG in plants (Blackburn and Gall 1978; Richards and Ausubel 1988; Meyne et al. 1989). Human telomeres are formed of degenerate variant repeats within the first 1 kilobase (kb) of the telomere sequence, with the most common being: TTGGGG, TCAGGG, CTAGGG and GTAGGG. A series of TTAGGG repeats is then found at the distal end and this sequence pattern varies between chromosomes allowing for the dissociation of alleles (Allshire et al. 1989; Baird et al. 1995).

Telomere repeats are also found at non-telomeric loci within chromosome arms or close to centromeres, and, these interstitial telomere sequences (ITSs) are thought to originate from fusion events which occurred during evolution (Meyne et al. 1990; Bolzan 2017; Baird 2018). Telomeres represent fragile sites within our genome and therefore, ITSs are thought to increase genomic instability as they present breakage sites within chromosome arms (Slijepcevic et al. 1996; Sfeir et al. 2009; Aksenova et al. 2013).

1.1.2 Telomere length

Variation in telomere length is observed between species with laboratory mice telomeres reaching up to 150 kb whilst human telomeres average between 5 and 15 kb and yeast cells

display telomeres of just 300 base pairs (bps) (Kipling and Cooke 1990; Louis and Borts 1995). It is also clear that telomere length determined at birth varies between individuals from the same species and is a genetically heritable trait (Slagboom et al. 1994; Okuda et al. 2002; Aviv et al. 2011). Humans have a longer lifespan than mice despite having shorter telomeres thus signifying that starting telomere length is not predictive of lifespan. Indeed, larger organisms undergo more cellular division and are therefore more prone to mutations and have thus evolved to have stringent tumour suppressive mechanisms, such as replicative senescence discussed in section 1.2.1, to limit large-scale genomic instability associated with critically short telomeres (Seluanov et al. 2008; Gomes et al. 2011). In addition, a single stranded overhang is present at the end of telomeres, not included in the overall telomere length, and varies in size, from as short as 24 nucleotides to over 400 nucleotides with the average measured around 100 nucleotides (Makarov et al. 1997; Cimino-Reale et al. 2001). The function of this structural property will be discussed in section 1.1.4.1.

1.1.3 Telomere-associated proteins

Numerous proteins reside at telomeres with some exclusively functioning at those sites, such as the Shelterin complex and telomerase, whilst others play a role at telomeres but also relocate to other parts of the genome when required, such as DNA damage sensing and repair proteins.

1.1.3.1 The Shelterin complex

The Shelterin complex is formed of six proteins which reside at telomeres (POT1, TPP1, TRF1, TRF2, RAP1 and TIN2) (de Lange 2005). TRF1 was the first subunit to be found through its ability to bind the double stranded DNA (dsDNA) alongside TRF2, its paralog (Zhong et al. 1992; Bilaud et al. 1997). Both proteins have two DNA Myb binding domains that recognise TTAGGG repeats exclusively and anchor the complex to the telomeric DNA, as well as protect the natural ends of chromosomes from the repair machinery (Broccoli et al. 1997; Court et al. 2005). Next, RAP1, binds TRF2 whilst TIN2 forms a bridge between TRF1 and TRF2 and also binds to TPP1 which stabilises the complex (O'Connor et al. 2006; Rai et al. 2016). Finally, POT1 binds the single stranded DNA (ssDNA) overhang present at the 3’ end of the telomere (Baumann and Cech 2001). TIN2, TPP1 and POT1 have also been found to directly recruit telomerase to induce telomere lengthening and increase telomerase processivity (telomerase action will be discussed in section 1.1.3.2) (Frank et al. 2015; Pike et al. 2019).

There are numerous Shelterin complexes across the entire telomere length and their role as a whole will be detailed in section 1.1.4.1.

1.1.3.2 Telomerase

Telomerase was first described in 1985 in *Tetrahymena* for its property to extend telomeres (Greider and Blackburn 1985). Telomerase is a reverse transcriptase enzyme that adds canonical telomeric repeats *de novo* to the distal end of chromosomes to compensate for telomere loss during replication (detailed in section 1.1.4.2) (Morin 1989). It is formed of two subunits: an RNA template encoded by the *hTERC* gene, located on chromosome 3, and the catalytic subunit encoded by the *hTERT* gene, located on chromosome 5 (Feng et al. 1995; Cong et al. 1999). hTERC is continually transcribed and can be detected in all cells whilst hTERT is only expressed when required (Avilion et al. 1996; Cong et al. 2002). The hTERC mRNA sequence is complimentary to the TTAGGG repeats present at the distal ends of telomeres. This RNA binds to the active site of hTERT, which in turn binds to the 3' single stranded overhang, and serves as a template to synthesise a new DNA sequence (Cong et al. 2002). Telomerase is recruited to telomeres through TPP1/TIN2/POT1 interaction which promotes enzymatic activity (Nandakumar and Cech 2013; Pike et al. 2019).

Telomerase activity is detected during early development followed by a decrease in levels and the absence of detection in human somatic cells. Telomerase is however active in the germ line and in most stem cells, such as hematopoietic cells, during adulthood allowing indefinite cellular division (Broccoli et al. 1995; Morrison et al. 1996; Wright et al. 1996). Upon maturation of stem cells, telomerase levels drop to become undetectable, although low levels are detected in cells with a high turnover rate such as basal skin cells (Harle-Bachor and Boukamp 1996). In contrast, telomerase is active in mice somatic cells and the silencing of telomerase has been proven to correlate with body mass, thus, the absence of the enzyme in larger organisms is suggestive of a tumour suppressive mechanism (Seluanov et al. 2007).

1.1.3.3 DNA damage sensing

Numerous proteins, termed DNA damage sensors, are involved in the recognition of DNA damage sites. One complex involved in the recognition of double stranded breaks (DSBs) is the MRE11-RAD50-NBS1 (MRN) complex. By definition, DSBs require both strands to be cut making these lesions lethal if unrepaired. The MRN complex recognises damage sites and activates transducers, such as ATM or ATR, which in turn phosphorylate cell cycle checkpoint proteins arresting the cell cycle, as well as the histone H2AX which signals for and recruits

appropriate downstream repair proteins (Uziel et al. 2003; Lamarche et al. 2010; Duursma et al. 2013). Interestingly, this complex also interacts with telomere-associated proteins, such as TRF2, therefore implying it plays an additional role in telomere maintenance and protection (Zhu et al. 2000). It has also been shown that the MRN complex may be involved in the recruitment of telomerase in immortal cells and loss of this complex resulted in a shortened G-rich overhang, required for telomerase anchoring and extension (Chai et al. 2006).

1.1.3.4 DNA repair

Many proteins involved in various DNA repair pathways are present at telomeres, notably key proteins involved in the repair of DSBs. Indeed, the two primary pathways utilised for the repair of such lesions are non-homologous end joining (NHEJ) alongside homologous recombination (HR). Ku is a subunit of the DNA-PK complex, involved in NHEJ-mediated repair, which binds to dsDNA to provide an anchoring point for other proteins involved in this pathway to ensure the successful repair of a DSB (Davis and Chen 2013). Interestingly, it also plays a major role in the protection of telomeres and has been found to interact with TRF1 and TRF2 to prevent NHEJ action to inhibit aberrant end-to-end fusions, contrary to the role it plays in repair (Hsu et al. 2000; Ribes-Zamora et al. 2013). The other mechanism used in the repair of DSBs is HR, and RAD51 is responsible for binding to the resected ssDNA and subsequent strand invasion into the homologous template DNA (Sung and Robberson 1995). By this means, RAD51 is also required at telomeres after replication to promote the invasion of the single stranded overhang present at telomeres into the dsDNA to enable the protection of telomeres (Verdun and Karlseder 2006), a role that is discussed in section 1.1.4.1.

1.1.4 Telomere function

1.1.4.1 Protection of natural chromosome ends

The Shelterin complex plays a key role in the protection of linear eukaryotic DNA. POT1, which binds to the ssDNA, alongside TRF2, which modulates the DNA to form a loop, enable the single stranded overhang to fold back on itself, termed telomeric or t-loop, and enables the ssDNA to insert itself into dsDNA forming a DNA triplex, also called displacement or d-loop (figure 1.1 A) (Griffith et al. 1999; Stansel et al. 2001). This mechanism prevents the natural ends of chromosomes from being detected by the DNA repair machinery, and

therefore mistaken for DSBs, and subsequently repaired through NHEJ or HR which in turn causes chromosomal rearrangements (Greider 1991). TRF2 is the main preventer of fusion events as it directly inhibits ATM migration to the damage site which subsequently prevents the downstream phosphorylation of H2AX and recruitment of repair proteins (van Steensel et al. 1998; Okamoto et al. 2013). This is confirmed by the fact that loss of TRF2 results in a high incidence of end-to-end fusions within 48 hours (Celli and de Lange 2005). This unique loop formation therefore protects the structural integrity of the genome.

1.1.4.2 Protection of protein coding DNA

1.1.4.2.1 The end-replication problem

DNA replication occurs during the S phase of the cell cycle and is the process by which DNA is duplicated in a semi-conservative fashion, by using the existing DNA as template for synthesis, prior to cell division to ensure DNA content is separated equally between daughter cells (Meselson and Stahl 1958). It occurs in the 3' to 5' orientation, therefore, the leading strand is synthesised in the same direction of the unwinding, whilst the lagging strand synthesis occurs in a discontinuous manor (figure 1.1 B). Firstly, RNA primers are added to both strands to allow the anchoring of DNA polymerase and the initiation of DNA synthesis. On the lagging strand, RNA primers are added as the DNA unwinds and DNA polymerase synthesises the DNA between primers forming Okazaki fragments, which are 100-200 nucleotides in length (Smith and Whitehouse 2012). Upon removal of these primers, DNA ligase fills in the gaps, but is unable to do so at the distal end of the chromosome. It is not clear where the most distal RNA primer is on the lagging strand and this therefore means a loss of 10-20 nucleotides (length of an RNA primer) up to 200 nucleotides (length of an Okazaki fragment) is incurred during this process whilst the leading strand is synthesised in full, also known as the end-replication problem (Ogawa and Okazaki 1980; Dai et al. 2009; Chow et al. 2012). This results in a gradual loss of telomeric sequence after every cell cycle.

1.1.4.2.2 Telomere erosion

Although only the lagging strand incurs a loss of nucleotides as a result of the end-replication problem, both strands shorten between each replication round by 50-200 nucleotides (Harley et al. 1990; Muraki et al. 2012). Leading and lagging strands appear to be processed differently with the removal of the final RNA primer on the lagging strand, which has been shown to be 70-100 nucleotides away from the end of the DNA, occurring soon after replication whilst a degradation at the 5' end on the leading strand appears later by EXO1

and APOLLO nuclease activity (Chow et al. 2012; Wu et al. 2012). The generation of a single strand 3' overhang, characteristic of the telomere structure, plays a fundamental role in the protection of telomeres to form the t-loop configuration as well as provide an anchoring point for telomerase and subsequent telomere extension by action of the enzyme (Lingner and Cech 1996; Makarov et al. 1997; Wright et al. 1997; Wu et al. 2010) (figure 1.1 C). Evidence that cell types divide at different rates and telomere erosion is dependent on the turnover has been established. In example, cardiac muscle cells undergo no to very little telomere erosion whilst cells originating from hematopoietic stem cells, such as T-cells and lymphocytes, undergo rapid telomere attrition, especially during childhood consistent with a high turnover rate at that time (Rufer et al. 1999; Bergmann et al. 2015; Sharifi-Sanjani et al. 2017).

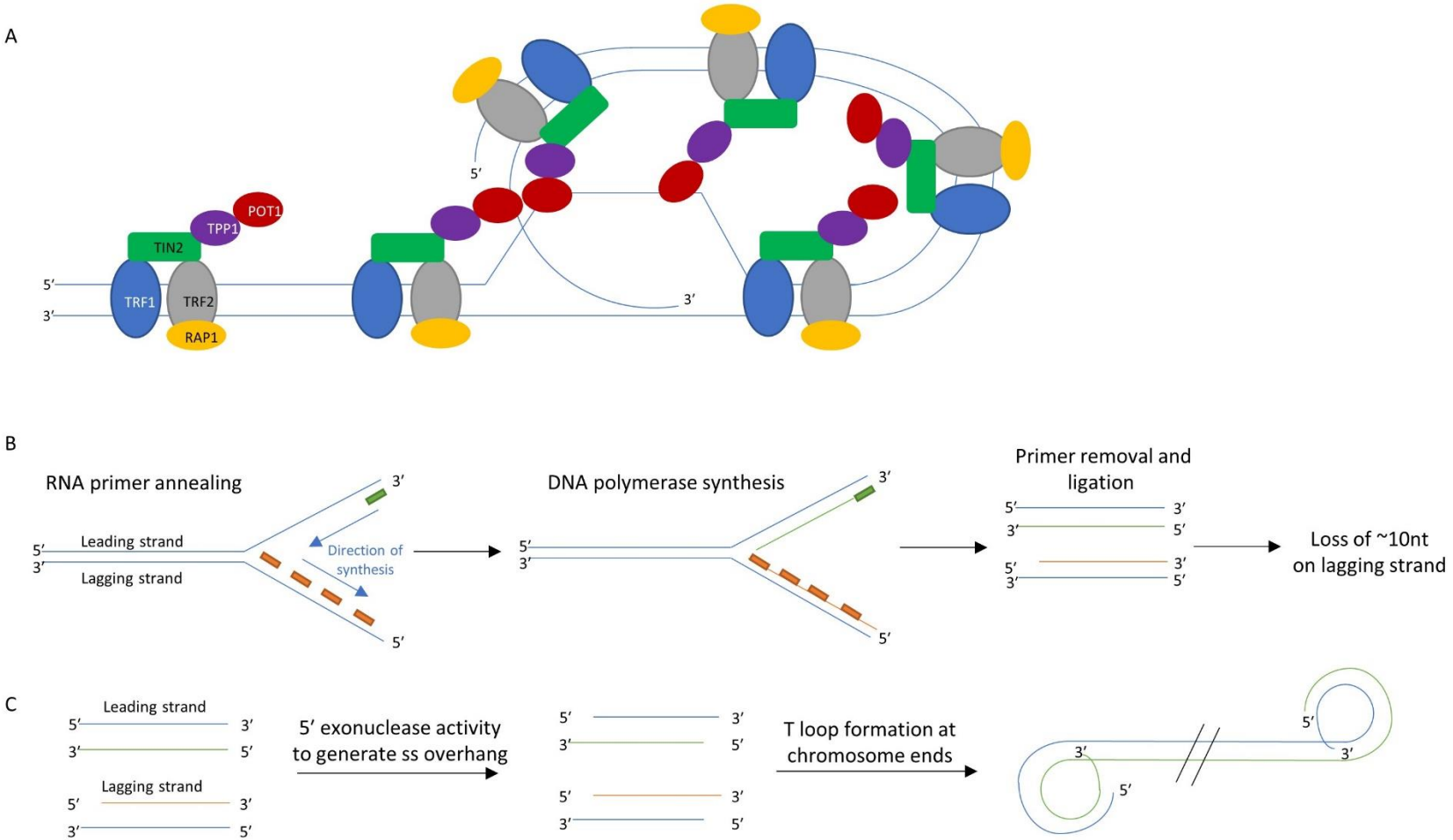


Figure 1.1: Schematic representation of telomere function.
A) The Shelterin complex bound to the telomere enabling the formation of a t-loop. B) DNA replication of the leading and lagging strands. C) Telomere erosion and generation of ssDNA overhang to enable loop formation. Adapted from de Lange 2005; Sampathi and Chai 2011.

1.1.5 Chromatin structure

1.1.5.1 Chromatin assembly

During interphase, the DNA is compacted into chromosomes encapsulated within the nucleus. The DNA is tightly bound to a nucleosome which is formed of 8 histones: $(H2A/H2B)_2$ and $(H3/H4)_2$ thus enabling the coiling of 145-147 bps (Luger et al. 1997). Histone variants exist for all histones apart from H4 in humans (Henikoff and Smith 2015). Notably, histone H3 is known to have five variants in humans: two canonical variants H3.1 and H3.2, for which expression is recorded at the start of S phase to ensure sufficient production of H3 histone for replication; and three replacement variants H3.3, CENP-A and H3.1t, which are expressed throughout the cell cycle (Hake and Allis 2006; Loyola and Almouzni 2007). Histone H3.3 varies by 5 and 4 amino acids from H3.1 and H3.2 respectively however, despite having similar sequences, these changes are thought to account for the difference in the role they each play (Goldberg et al. 2010). The role of histone H3.3 will be discussed in section 1.4.3.

Histones undergo post-translational modifications ranging from acetylation and phosphorylation to methylation which enables the interaction with numerous proteins such as histone chaperones and chromatin remodelers (Allfrey et al. 1964; Bannister and Kouzarides 2011). Indeed, histones can be incorporated at particular regions of the genome, as well as at specific times in the cell cycle, to ensure chromatin stability and condensation, such as H3.3 is incorporation at telomeres following tri-methylation of lysine 9, or to provide access to a gene for expression, such as incorporation of histone H3.3 within promoters of transcribed genes (Chow et al. 2005; Goldberg et al. 2010; Lewis et al. 2010). Meanwhile, H2AX phosphorylation triggers the DSB repair machinery thus showing the extensive role histone modifications play in cellular processes (Rogakou et al. 1998).

1.1.5.2 Telomere transcription

Despite the telomere sequence not encoding for proteins and constituting repetitive DNA, this area of the genome is not silent. Indeed, telomeres have been shown to be actively transcribed from the sub-telomere DNA, containing CpG islands serving as promoters for transcription, to the telomere repeat array into telomeric repeat containing RNA (TERRA) by RNA polymerase II, using the C-rich strand as a template (Azzalin et al. 2007; Schoeftner and Blasco 2008; Nergadze et al. 2009). By this means, only a subset of telomeres are known to be transcribed, including 5p, 9p, 17p or 21q amongst others (Nergadze et al. 2009). TERRAs localise at all telomeres post-transcriptionally, but not exclusively, and are thought to play

an important role in the protection of telomeres from the repair machinery, as depletion of TERRA induces telomere dysfunction (Azzalin et al. 2007; Lopez de Silanes et al. 2014; Chu et al. 2017). Similarly, an increase in TERRA levels is also associated with dysfunctional telomeres and diseases such as cancer inducing rapid telomere shortening and therefore, a tight control of TERRA is required for appropriate function (Maicher et al. 2012a; Maicher et al. 2012b).

TERRAs can bind DNA to form DNA-RNA hybrids termed R-loops at the time of transcription, when DNA is unwound and accessible, which results in a DNA-RNA duplex and an ssDNA strand prone to DNA-damaging agents; their exact role remains unclear (Thomas et al. 1976; Balk et al. 2013). Interestingly, these secondary structures appear to be more stable than the native dsDNA, although, an increase in these R-loops disrupts homeostasis and is associated with high levels of genomic instability and therefore diseases such as cancer (Roberts and Crothers 1992; Richard and Manley 2017). Indeed, R-loops are thought to affect DNA replication as they form a barrier for polymerase resulting in replication fork stalling and collapse (Gan et al. 2011). In addition, they can also inhibit gene expression by downregulating methylation of CpG islands found upstream of a subset of genes and acting as promotors (Ginno et al. 2012). These structures can however be enzymatically resolved, notably with RNase H1 (Parajuli et al. 2017).

1.2 Telomere-associated ageing and diseases

1.2.1 Senescence

Replicative senescence is an irreversible process by which cells stop dividing whilst remaining active and influencing the microenvironment (van Deursen 2014). Senescence can be triggered by numerous processes which cause stress such as damage induced by reactive oxygen species (ROS) or telomere dysfunction (von Zglinicki 2002; Deng et al. 2008). Primary somatic cells that undergo cellular division have a finite replicative lifespan and after 60 population doublings (PDs) on average, also known as the Hayflick limit, they stop dividing (Hayflick and Moorhead 1961). Upon loss of telomeric chromatin as a result of replication, senescence is initiated by the accumulation of short telomeres which in turn trigger a TP53 dependent G1/S cell cycle arrest, due to telomeres being too short to fulfil their protective role, resulting in the loss of the t-loop formation and the exposure of the natural ends of chromosomes to the DNA damage response, resulting in a cell-cycle checkpoint response (Shay et al. 1991; von Zglinicki et al. 2005). In addition, phenotypic changes are observed in senescent cells as the chromatin structure alters and gene expression patterns adapt to release pro-inflammatory cytokines, chemokines, growth factors, etc, also known as the senescence-associated secretory phenotype (SASP), which influences the surrounding environment (Acosta et al. 2013; Sun et al. 2018). Therefore, senescent cells accumulate over time in tissues and senescence can be triggered in neighbouring cells by the release of these pro-inflammatory particles (Nelson et al. 2012). This in turn revealed the use of acute senescence for tissue repair or development for instance. The phenotype is thought to be transient in these cases to enable cellular processes to take place and limit fibrosis in the case of repair (Krizhanovsky et al. 2008; Jun and Lau 2010; Rajagopalan and Long 2012). Senescence is therefore a stringent tumour suppressive mechanism, preventing cells with dysfunctional telomeres from dividing further thus inhibiting genomic rearrangements, and is an evolutionary trait of organisms which have longer lifespans which present the potential of accumulating more mutations over time (Sager 1991; Deng et al. 2008).

1.2.2 Telomere erosion and diseases

In normal individuals, telomere lengths at birth are similar in all tissues however the rate of telomere erosion has shown conflicting results (Youngren et al. 1998). It would be expected that cells that undergo limited proliferation, such as muscle cells, would have a slower rate

of telomere attrition than fast dividing cells, such as hematopoietic cells. However, Daniali et al showed a similar rate of erosion across both cell types over the course of a lifetime despite longer telomeres in muscle cells versus leukocytes (Daniali et al. 2013). They suggested that leukocytes undergo rapid telomere erosion early in life, followed by a stabilisation whilst muscle cells undergo a similar rate of erosion throughout the lifespan of the individual, thus also accounting for the differences in telomere length between the cell types (Daniali et al. 2013). In addition, telomeres measured under the threshold of 5 kb were seen to be predictive of imminent death and coincided with older individuals, 80 years of age and above in most cases (Steenstrup et al. 2017).

An extensive breadth of studies reports the impact of the environment on telomere biology and especially on the rate of erosion (Romano et al. 2013; Dugdale and Richardson 2018). The effects of smoking, obesity, and prolonged exposure to stress amongst others have been evaluated and all appear to impact telomere length and the rate of telomere attrition, showing an increased rate of telomere loss (von Zglinicki 2002; Song et al. 2010). Interestingly, exposure to these environmental factors in-utero, notably the mother's level of stress, has a greater effect on telomere length, as cells divide more rapidly during development, and can impact the rate of telomere attrition throughout the individuals lifetime (Entringer et al. 2011). As mentioned previously, the telomere length at birth varies amongst individuals, however, the starting telomere length rather than the rate of telomere erosion has been suggested to be predictive of developing conditions such as cardiovascular diseases (Toupance et al. 2017; Benetos et al. 2018; Benetos et al. 2019). The reason for this observation remains unclear but may be linked to causal events such as a higher rate of insulin seen in patients with cardiovascular disease, although this remains to be fully tested (Zhan et al. 2017). In contrast, despite observing shorter telomeres in patients with osteoarthritis, this is thought to be a consequence of accelerated telomere attrition, due to exposure to stress, and results in an increase in senescent cells contributing to the disease phenotype (Martin and Buckwalter 2001; Kuszel et al. 2015). Senescence is a tightly controlled phenomenon and aberrant rates of telomere erosion result in the accumulation of senescent cells and premature ageing (Childs et al. 2015). This is a common observation in genomic instability syndromes.

1.2.3 Genomic instability syndromes

Genomic instability syndromes include a wide range of diseases, with various symptoms and causes, but all exhibit forms of chromosomal instability. In example, Ataxia telangiectasia (AT) is a syndrome caused by a mutation in the ataxia telangiectasia mutated (*ATM*) gene which encodes the ATM protein, a DNA damage transducer, involved in DNA damage sensing and phosphorylation of the histone H2AX for recruitment of downstream repair (Savic et al. 2009; Rothblum-Oviatt et al. 2016). Patients are sensitive to ionising radiation and this therefore results in the accumulation of DNA damage and genomic instability (Painter and Young 1980). Some patients also have immunological abnormalities and together with the increased genomic instability, patients are more susceptible to developing cancers, especially lymphoma and leukaemia (Nowak-Wegrzyn et al. 2004; Reiman et al. 2011; Suarez et al. 2015).

There are many other syndromes that are regrouped in this category including two syndromes which are caused by mutations in RecQ helicases causing replication and DNA repair defects: Bloom syndrome, caused by a mutation in the BLM protein, and Werner syndrome, caused by a mutation in the WRN protein (Burla et al. 2018). The WRN protein has been shown to unwind areas of the genome that are difficult to replicate, such as telomeres, and ensures appropriate lagging strand synthesis (Crabbe et al. 2004). The BLM protein interacts with the Shelterin component TRF1 and appears to play a role in telomere maintenance and, also minimises the rate of crossovers associated with HR during the repair of DNA lesions (Karow et al. 2000; Barefield and Karlseder 2012). Loss of BLM or WRN results in telomere dysfunction through end-to-end fusions and high rates of recombination, notably sister-chromatid exchanges (SCE), and an accelerated rate of telomere erosion which in turn leads to premature ageing (Chaganti et al. 1974; Faragher et al. 1993; Wyllie et al. 2000; Bohr 2002). Patients with Werner syndrome present a higher risk of developing sarcomas whilst Bloom syndrome patients are susceptible to any malignancy due to the accumulation of genomic instability in the absence of these proteins (Oshima 2000; Callen and Surralles 2004).

1.2.4 Crisis and cancer

If telomeres continue to shorten as a function of cell division, beyond senescence, due to cell cycle checkpoint proteins, such as TP53 or Rb, being mutated, short telomeres begin to become dysfunctional. This therefore leads to end-to-end fusions, including sister-chromatid

and inter-chromosomal fusions, dicentric chromosomes and overall large-scale chromosomal rearrangements as telomeres are now recognised as DSBs (Counter et al. 1992; Capper et al. 2007). Cells continue to divide despite extensive damage resulting in rearranged karyotypes due to tearing of chromosomes during mitosis. This phase is known as crisis, during which most cells will die due to extensive lethal mutations however, a rare proportion of cells will survive leading to clonal malignant growth (figure 1.2) (von Morgen and Maciejowski 2018). Cells achieve replicative immortality by maintaining telomere length by one of two known mechanisms: 85-90% of cancers upregulate the enzyme telomerase whilst the remaining 10-15% with activate the alternative lengthening of telomeres (ALT) (Heaphy et al. 2011b).

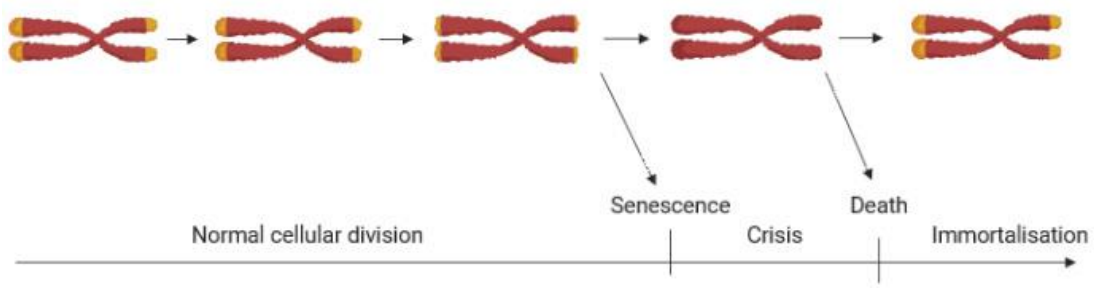


Figure 1.2: Schematic representation of telomere erosion as a function of cellular division.
Telomere erosion triggers replicative senescence or crisis resulting in cell death or immortalisation through maintenance of telomere length in cancer. Figure generated using BioRender.

As previously mentioned, telomerase is present in the germline and stem cells but not in somatic cells outside of development. The mechanism of telomerase-mediated telomere elongation is well understood and its upregulation in cancer has been well characterised. hTERT expression is driven by multiple changes such as an amplification of the *hTERT* gene or, mutations within the *hTERT* promoter thus allowing transcription factors to bind and drive expression of the gene (Zhang et al. 2000; Huang et al. 2013). A five-fold increase in hTERC levels has been observed in some cancer cells, and appears to depend on the successful upregulation of hTERT combined with an increase in the rate of transcription of the *hTERC* gene (Yi et al. 1999). In contrast, the ALT mechanism is still poorly understood, and the initiation and maintenance have yet to be fully elucidated.

1.3 Alternative lengthening of telomeres

The ALT mechanism was first described in 1995 by Tracy Bryan who uncovered the unique property of telomere elongation and therefore replicative immortality in cells that lack telomerase activity (Bryan et al. 1995). The mechanism of telomere elongation in the context of ALT varies greatly from telomerase upregulation which adds telomeric repeats *de novo*. ALT telomeres undergo rapid extension through homology directed repair (HDR), notably break-induced replication (BIR) (Dilley et al. 2016; Roumelioti et al. 2016). The initiation and maintenance of this mechanism remain largely unknown, although, various hallmarks associated with ALT and which contribute to the phenotype have been uncovered in the last 25 years. These include extrachromosomal DNA present in the cytoplasm of cells termed C-circles, or ALT-associated PML bodies (APBs) which provide essential proteins required for HDR-mediated elongation of telomeres, alongside numerous other pathways discussed below which contribute to the ALT phenotype and show variability amongst cancer subtypes (Yeager et al. 1999; Henson et al. 2009; Sommer and Royle 2020).

1.3.1 Mechanism of telomere length maintenance

1.3.1.1 Normal HDR repair

HDR is a faithful repair pathway primarily used to resolve DSBs alongside the NHEJ pathway. HDR is also involved in the resolution of replication fork stalling or collapse (Costes and Lambert 2012; Wilhelm et al. 2016). There are four main types of HDR: single strand annealing (SSA), double strand break repair (DSBR), synthesis-dependent strand annealing (SDSA) and break-induced replication (BIR) (Heyer et al. 2010). In the event of the repair of a DSB through HDR, DSBR and SDSA are primarily used and begin with EXO1-mediated resection of DNA in the 5’ to 3’ direction, either side of the break, to generate single stranded overhangs to which RPA binds, to protect the exposed fragile ssDNA and to recruit downstream proteins (Ruff et al. 2016). RPA is then displaced from ssDNA by various proteins such as RAD52 and BRCA2 to enable the binding of RAD51, a step that is essential in strand invasion of the homologous template (Sung and Robberson 1995; Sugiyama and Kowalczykowski 2002; Jensen et al. 2010). In the case of SDSA, one strand invades the homologous template, temporarily forming a d-loop, and allowing for the synthesis of the DNA using the template homologous strand. The d-loop is then resolved through the dissociation of the newly synthesised strand and the gaps are filled through ligation action

(figure 1.3) (Nassif et al. 1994). In contrast, DSBR is a complex pathway during which the four DNA strands are joined to form a double Holliday junction, resulting in complex exchanges of DNA and inducing heterogeneity within the sequence upon resolution of the structure (figure 1.3) (Holliday 2007; Wyatt and West 2014). Crossovers are however rare in somatic cells, and occur more frequently during meiosis to ensure genetic diversity (Lam and Keeney 2014).

In the case of single strand annealing (SSA), which occurs when a DSB arises between homologous repeat sequences, DNA is similarly resected in the 5' to 3' direction and the repeat sequences are annealed through RAD52 action, followed by a removal of the single stranded overhangs by ERCC1 (figure 1.3) (Van Dyck et al. 2001; Motycka et al. 2004). SSA is therefore associated with deletions at the site of repair (Stark et al. 2004). Finally, BIR is primarily used to repair replication stress and usually involves a one ended break, such as the natural ends of chromosomes. This pathway will be discussed in more detail in section 1.3.1.2 and figure 1.4 as it is thought to be the mechanism used to elongate telomeres in the context of ALT upregulation for survival in cancer.

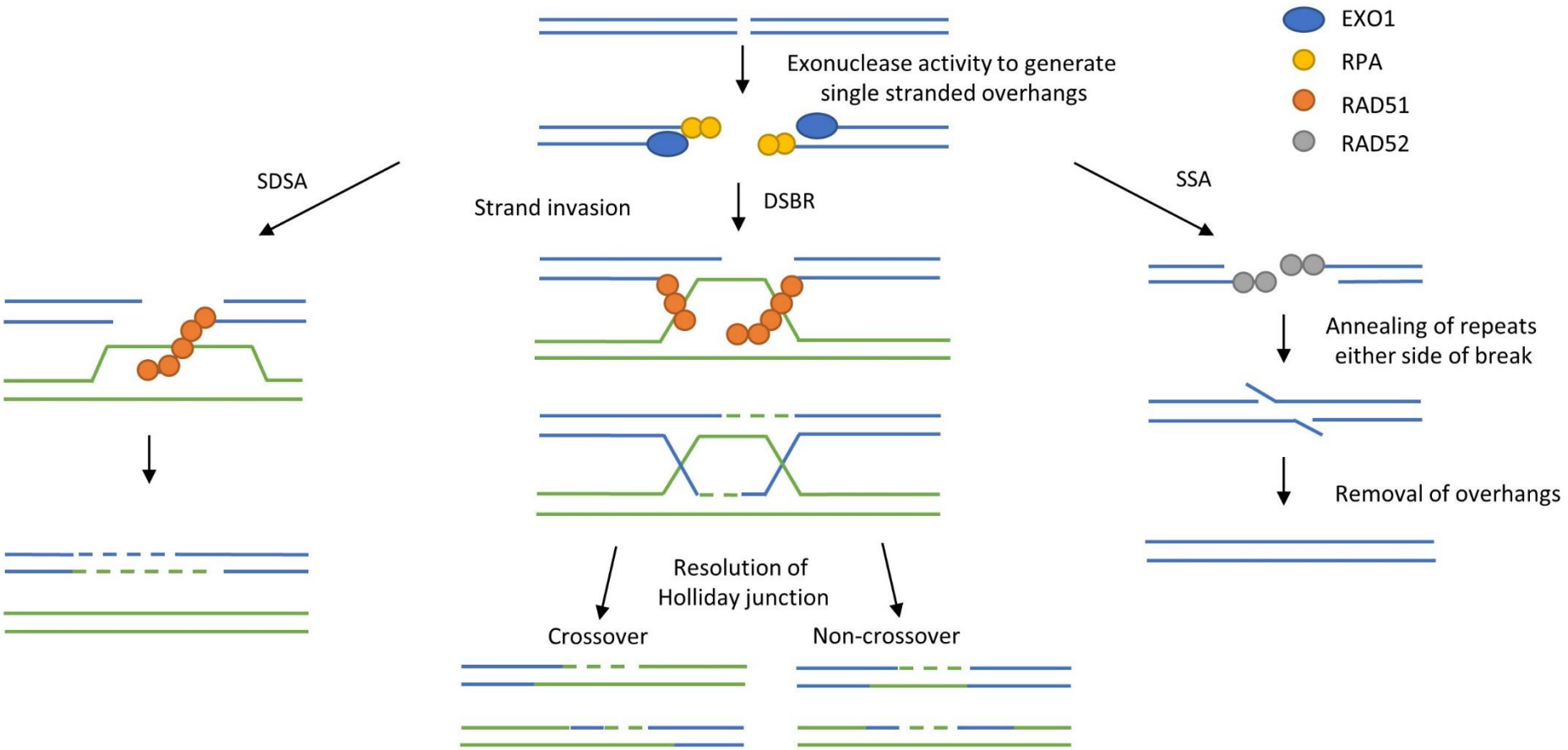


Figure 1.3: Repair of DSBs through HDR.
DSBs can be repaired through SDSA, DSBR and SSA resulting in accurate resolution of breaks with varying outcome of heterogeneity in DNA sequences dependent on the pathway utilised. Adapted from Heyer et al. 2010.

1.3.1.2 BIR-mediated telomere elongation in ALT positive cells

The exact mechanism of elongation and maintenance of telomeres in the context of ALT is still unclear, however, HDR has been strongly linked to the ALT mechanism. For instance, evidence that tags inserted within the telomere sequence were exchanged between sister chromatids as well as between chromosomes was established (figure 1.4 A and B) (Dunham et al. 2000). It is also thought that the length heterogeneity as well as complex mutations within the proximal region to the sub-telomeres observed in ALT cells results from inter-chromosomal exchanges (Varley et al. 2002; Liu et al. 2018). The WRN protein appears to play a role in the ALT mechanism as it further promotes inter-allelic recombination, perhaps due to its ability to unwind complex structures such as t-loops and G quadruplexes (G4) at telomeres (Mendez-Bermudez et al. 2012). In addition, numerous HDR-related proteins such as RAD51 and RAD52 were isolated in APBs, markers of ALT activity, thus further suggesting the involvement of recombination in ALT positive cells (Yeager et al. 1999).

BIR has emerged as the leading HDR pathway in the context of ALT for maintenance of telomere length as it involves a one-ended repair (Figure 1.4 B) (Dilley et al. 2016; Roumelioti et al. 2016). The first step of BIR involves the resection of DNA to form single stranded overhangs, similar to other HDR pathways. From there, two BIR pathways have been identified: RAD51-dependent and -independent (Sakofsky and Malkova 2017). When RAD51 is involved, it binds to the single stranded overhang and invades the homologous dsDNA to initiate synthesis (Davis and Symington 2004). The RAD51-independent pathway relies on RAD52 for annealing of ssDNA in a similar fashion to SSA (Malkova et al. 1996). ALT cells appear to use both mechanisms, however, the initial elongation event is thought to arise from using C-circle DNA, a hallmark of ALT activation, as template for RAD51-independent BIR, allowing for the annealing of the ssDNA to the telomere and proceeding to a rolling-circle amplification (RCA) for elongation (figure 1.4 C) (McEachern and Blackburn 1996; Tomaska et al. 2009). C-circles are thought to arise following replication fork collapse resolution or RCA of telomeres and is further discussed in section 1.3.5.4 (Zhang et al. 2019b). Linear telomeric DNA found in APBs, another hallmark of ALT, was also speculated to being used as a template for elongation, although short in length thus not suggestive of rapid and significant elongation observed in ALT cells (Yeager et al. 1999; Henson et al. 2002). This is then followed by RAD51-dependent BIR between telomeres for continual telomere elongation (Natarajan and McEachern 2002; Zhang et al. 2019a).

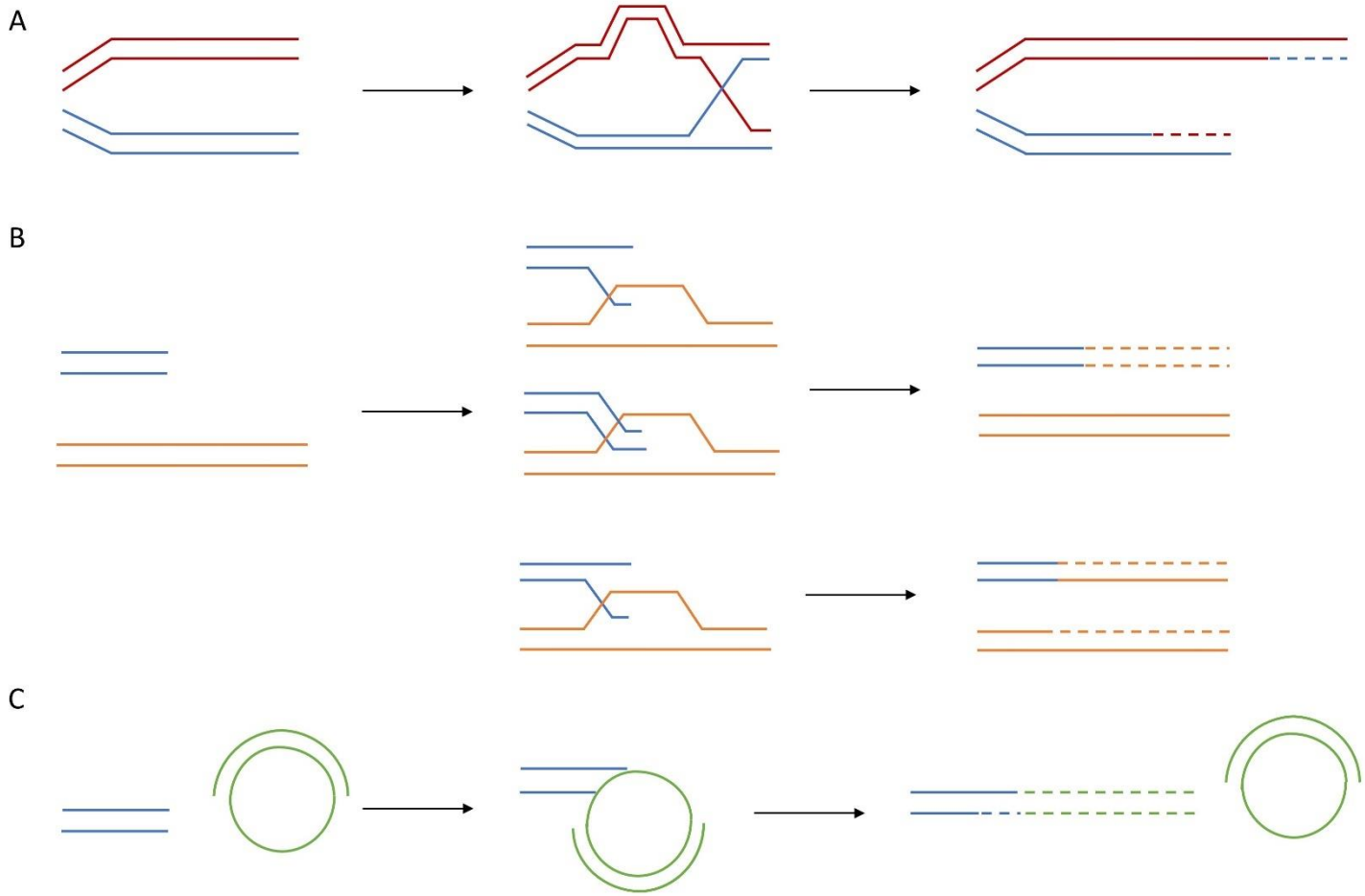


Figure 1.4: Putative HDR-mediated telomere elongation in the context of ALT.
A) Unequal sister-chromatid exchange. B) Break-induced replication with strand invasion of the short telomere into the long telomere for extension with two possible outcomes. C) Rolling circle amplification using partially single stranded C-circle as template for telomere elongation. Adapted from Durant 2012; Sakofsky and Malkova 2017.

1.3.2 Use of ALT in various organisms

The initiation and maintenance of the ALT mechanism remains largely unclear although this phenomenon has been seen to occur in various organisms, spontaneously or artificially induced, and provide strong models for a better understanding of this mechanism.

1.3.2.1 Yeast

In *S. cerevisiae* two telomerase-independent types of survivors emerged spontaneously from cultures (Lundblad and Blackburn 1993; Teng and Zakian 1999). To study this phenomenon, the *est1* gene, which encodes for telomerase in yeast, was knocked out and initially no difference in viability was observed. Telomeres eroded until reaching senescence at which point an increase in cell death was observed whilst a subset of cells bypassed senescence (Lundblad and Szostak 1989; Lundblad and Blackburn 1993). Yeast telomeres have distinctive X and Y elements, which resemble sub-telomere sequences, interspersed with telomere repeat tracts (Chan and Tye 1983; Shampay et al. 1984). Type I survivors show an amplification of the X and/or Y elements and depend on RAD52 and RAD51. In contrast, type II survivors show an increase in telomere repeats and present ALT-like features and depend on RAD52 and RAD50 (Teng and Zakian 1999; Teng et al. 2000; Chen et al. 2001). In type II survivors, telomeres were shown to erode until the sudden appearance of long telomeres containing a high proportion of telomeric repeats thus suggesting the use of telomeres as templates for elongation, and also evidence that only short telomeres are elongated for survival (Teng et al. 2000). In addition, they are dependent on HDR-related proteins, thus further confirming the use of recombination for elongation in these cells consistent with ALT upregulation (Chen et al. 2001). It is however important to note that this spontaneous phenomenon is rare (less than 10%), as it has been shown that these organisms adapt to low-levels of telomerase expression by the loss of a copy of chromosome VIII whilst still conferring replicative immortality and only a complete loss of telomerase results in the ALT phenotype (Teng et al. 2000; Millet et al. 2015).

1.3.2.2 C. elegans

To investigate the activation process associated with ALT, the telomerase coding gene, *trt-1*, was knocked out in *C. elegans* populations to induce telomere erosion and chromosomal instability once telomeres reached a critically short length (Lackner et al. 2012). A small percentage of nematodes survived beyond 200 generations, called survivors, and appeared to have adapted to the loss of telomerase and to the alterations occurred during ALT

upregulation. They showed evidence of the presence of hallmarks of ALT, discussed in section 1.3.5, such as C-circles and heterogeneous telomere lengths as well as chromosomal end-to-end fusions which presumably resulted from a crisis phase these cells had undergone (Cheung et al. 2004; Cheng et al. 2012). Deletion of the *C. elegans* equivalent of the human Shelterin component POT1 protein (CeOB1 and 2) showed an increase in telomere elongation and heterogeneity, resembling an ALT phenotype. Double knockout of POT1 and trt-1 facilitates the survival therefore suggesting a potential role of POT-1 in the ALT mechanism (Raices et al. 2008; Lackner and Karlseder 2013).

1.3.2.3 Mammals

Until recent years, ALT did not appear to be upregulated in normal human cells, unlike telomerase which is active in the germline and stem cells, and was solely associated with cancer. It has now been shown that ALT can be upregulated transiently to compensate for damage incurred at telomeres by chronic oxidative stress or radiation (Berardinelli et al. 2010; Coluzzi et al. 2017; De Vitis et al. 2019). In addition, ALT activity was observed in normal mouse somatic cells, but not in the germline, thus suggestive of a background of ALT activity (Neumann et al. 2013). Therefore, it cannot be excluded that low levels of ALT activity, below detection thresholds, may occur in normal human cells and that both telomerase and ALT are activated but below the detection level. In this situation, upon induction of a telomere crisis, one mechanism may outcompete the other to provide replicative immortality. It is however important to note that normal and telomerase positive cells have been shown to repress the ALT mechanism and it has been suggested that ALT may occur due to a repression of telomerase activity (Perrem et al. 1999). Indeed, as shown in other organisms above, survival in the absence of telomerase in *S. cerevisiae* and *C. elegans* showed an upregulation of ALT, however even low levels of telomerase were enough to inhibit ALT in yeast (Millet et al. 2015). In addition, telomerase activity is regarded an evolutionary trait in higher order eukaryotes thus providing a novel mechanisms for survival and could therefore explain the higher propensity to activate telomerase for survival in human cancers (de Lange 2004).

1.3.3 Prevalence in human cancers

Heaphy et al conducted a large study identifying the telomere maintenance mechanism (TMM) used in over 6,100 tumour samples from over 150 cancer subtypes combining a range of sarcomas, carcinomas and lymphomas (Heaphy et al. 2011b). From this study, carcinoma

samples from a broad variety of tissues throughout the body including breast, lung, skin, CNS, bladder, were evaluated to confirm telomerase upregulation, or ALT activation, for survival. Only 0.9% (41) were positive for ALT whilst the rest upregulated telomerase indicating that epithelial cells preferentially upregulate telomerase for survival. Epithelial cells originate from various stem cells, which show telomerase activity during development (Wright et al. 1996). Telomerase is then silenced upon maturation and differentiation of cells perhaps pre-disposing these cells to re-activating telomerase for survival (Broccoli et al. 1995). Nonetheless, a subtype of breast cancer, ductal carcinoma, showed a small percentage of ALT positivity (3/71; 4%) and coincided with a HER-2 overexpression which was associated with a poorer prognosis (Subhawong et al. 2009).

In contrast, the percentage of ALT positivity in sarcomas is greatly increased implying that cells of mesenchymal origin preferentially upregulate the ALT mechanism for survival (Heaphy et al. 2011b; Xiao et al. 2013). The reason for this is unclear, although, mesenchymal stem cells (MSCs) use little to no telomerase during development, unlike other stem cells, and could therefore be pre-disposed to upregulating ALT (Parsch et al. 2004). Combining several studies which address the ALT status of various sarcoma subtypes revealed an overall 40% positivity of ALT in leiomyosarcomas, liposarcomas, osteosarcomas and malignant fibrous histiocytomas (MFH) (table 1.1). A variation among subtypes of sarcomas is also observed with the highest rate of 56% overall in leiomyosarcomas down to 24% in liposarcomas. Although all studies determined ALT positivity by APB signal through fluorescent in situ hybridisation (FISH) combined with telomere length, a variation in thresholds used may explain the differences between studies. Furthermore, evidence of liposarcomas without APBs nor telomerase, but with other ALT-associated changes, such as recombination mediated mutations, have been identified whilst another study showed no distinct TMM used in 50% of liposarcoma samples tested for telomerase and ALT. Therefore, further confirmation of ALT status testing for another hallmark using a more robust protocol, such as the C-circle assay, should be considered (Henson et al. 2005; Costa et al. 2006; Jeyapalan et al. 2008). In addition, some studies have highlighted the co-existence of ALT and telomerase in a subset of sarcomas although it is unclear whether these co-exist within the same population of cells, or if it is a result of two cell populations co-existing with separate TMMs and if this is maintained on the long-term (Costa et al. 2006; Matsuo et al. 2009). Nonetheless, these studies have enabled the research community to narrow down the spectrum of cells that upregulate ALT which is important for the development of diagnostic, therapeutic and prognostic markers.

Type	Number of ALT+	% of ALT+	Study
CARCINOMAS			
Breast ductal carcinoma	3/71	4%	(Subhawong et al. 2009)
	5/251	2%	(Heaphy et al. 2011b)
Total	**8/322**	**2.5%**	
All	41/4756	0.9%	(Heaphy et al. 2011b)
LYMPHOMAS			
All	0/104	0%	(Heaphy et al. 2011b)
SARCOMAS			
Leiomyosarcoma	31/59	53%	(Heaphy et al. 2011b)
	8/13	62%	(Henson et al. 2005)
	51/92	55%	(Liau et al. 2015b)
	33/54	61%	(Yang et al. 2015)
Total	**123/218**	**56%**	
Liposarcoma	36/139	26%	(Costa et al. 2006)
	9/38	24%	(Heaphy et al. 2011b)
	3/9	33%	(Henson et al. 2005)
	23/111	21%	(Lee et al. 2015a)
Total	**71/297**	**24%**	
Osteosarcoma	27/58	47%	(Henson et al. 2005)
	12/14	86%	(Chen et al. 2014)
Total	**39/72**	**54%**	
Malignant fibrous histiocytoma	17/22	77%	(Henson et al. 2005)
	14/43	33%	(Matsuo et al. 2009)
Total	**31/65**	**48%**	
All from above	264/652	40%	
OTHER			
Astrocytoma	53/214	25%	(Abedalthagafi et al. 2013)
	38/115	33%	(Heaphy et al. 2011b)
	17/50	34%	(Henson et al. 2005)
Total	**108/379**	**28%**	
Glioblastoma	35/202	17%	(Heaphy et al. 2011b)
	19/77	25%	(Hakin-Smith et al. 2003)
Total	**54/279**	**19%**	
Primary pancreatic neuroendocrine tumour	98/321	31%	(Singhi et al. 2017)
	59/269	22%	(Kim et al. 2017)
Total	**157/590**	**27%**	
All from above	319/1248	26%	

Table 1.1: Prevalence of the ALT mechanism for survival in carcinomas, lymphomas, sarcomas and other tumours.

1.3.4 Therapy and prognosis

No current diagnostic test or treatment protocols are used to target ALT specifically in cancers as the mechanism of initiation and maintenance remain largely unclear. Thus far, the

best line of treatment for ALT-positive tumours is surgery followed by post-operative radiotherapy and/or chemotherapy. ALT positive tumours present more complex karyotypes and poorer differentiation than ALT negative samples thus indicating a poorer prognosis for patients with tumours utilising ALT for survival (Liau et al. 2015a; Ren et al. 2018). A meta-analysis of several studies, which included primarily liposarcomas and leiomyosarcomas, was carried out to further assess the prognosis in ALT positive versus ALT negative patients. It showed that the risk of death was doubled in ALT positive patients, whilst independent studies showed similar outcomes in other soft tissue sarcomas (STS) (Matsuo et al. 2010; Liau et al. 2015b; Lawlor et al. 2019). In contrast, ALT positivity improved the overall survival in glioblastoma with approximately 66% of patients still alive after 1 year versus 28% after the same period for telomerase positive patients (Hakin-Smith et al. 2003; McDonald et al. 2010). Interestingly, ALT positivity in osteosarcoma does not affect prognosis and in fact, the lack of known TMM upregulation shows better survival in these patients (Ulaner et al. 2003). It is however not clear why such a stark difference is observed between cancer subtypes.

1.3.5 Hallmarks of ALT

The ALT pathway is associated with numerous hallmarks that have been found in the last 25 years and enable the differentiation from telomerase positive tumours. It is however not yet understood when these alterations occur and if a sequential order is required for successful survival through the ALT mechanism (Bryan et al. 1995; Henson et al. 2002).

1.3.5.1 Lack of telomerase

ALT was initially discovered in cancers that lacked telomerase activity despite their telomeres undergoing extension for survival (Bryan et al. 1995). Upon assessment of telomerase components, levels of hTERT mRNA were undetectable due to a hypermethylation of the gene promoter, whilst hTERC was found to be both wild-type (WT) or absent and exogenous expression of hTERT showed restoration of telomerase expression when WT hTERC was present (Kilian et al. 1997; Wen et al. 1998; Dessain et al. 2000). This suggests that telomerase activity is hindered in ALT positive cells by repressed hTERT expression although the reason for this is unknown (Henson et al. 2002; Lafferty-Whyte et al. 2009).

1.3.5.2 Telomere length heterogeneity

ALT cells present extremely heterogeneous telomere lengths ranging from 3 kb to 50 kb, in contrast to telomere lengths measured below 8 kb in normal or telomerase positive cells (Bryan et al. 1995; Henson et al. 2002; Baird et al. 2004). Work in yeast showed the erosion of telomeres followed by a sudden increase in length suggesting only short telomeres undergo ALT-like elongation (Teng et al. 2000). It is however not clear whether there is a consistent length to which telomeres need to erode to before being elongated or if indeed telomeres are extended to a specific length.

1.3.5.3 APBs

ALT-associated promyelocytic leukaemia bodies (APBs) are structures which contain the PML protein alongside telomeric DNA and telomere associated proteins such as TRF1 and TRF2, key components of the Shelterin complex, or RAD51 and RAD52, involved in the HDR pathway (Yeager et al. 1999; Wu et al. 2000; Wu et al. 2003). APBs are detected through FISH probing for telomere repeat DNA (TTAGGG) and the PML protein (Yeager et al. 1999). This is the most common method used in the literature to evaluate the ALT status in tumour samples as the majority of ALT positive cancers are positive for APBs. In addition, APBs are readily detected during late S/G2 and M phases of the cell cycle (Grobelny et al. 2000). Their role in ALT is unclear although they are thought to localise at telomeres and provide HDR-related proteins as well as telomere DNA to be used as template to these sites for elongation thus suggesting their involvement in telomere maintenance (Draskovic et al. 2009). Indeed, RAD51, RAD52 and RPA are commonly found in APBs and are required in BIR-mediated repair and elongation (Yeager et al. 1999). They also appear to be a requirement for ALT maintenance as repression of the ALT mechanism is associated with a loss of APB signal and formation (Perrem et al. 1999; Jiang et al. 2007).

1.3.5.4 Circular extrachromosomal terminal repeats

C-circles are partially single stranded circular extrachromosomal terminal repeat (ECTRs) DNA present in the cytoplasm of ALT positive cells. As the name suggests, these are C-rich sequences which can be amplified through RCA and radioactive labelling and subsequently quantified according to intensity (Henson et al. 2009). The quantity of C-circles is thought to be directly correlated with levels of ALT activity and once more, these structures disappear rapidly upon ALT repression suggesting their involvement in ALT maintenance (Henson et al. 2009). The reason for the accumulation of these ECTRs is unclear or the role they play in ALT

upregulation and maintenance however, it is thought that these are used as templates for the initial elongation of telomeres through RCA, which has the potential for generating very long molecules and could account for the sudden increase in overall telomere length seen in ALT cells (McEachern and Blackburn 1996; Henson et al. 2009; Tomaska et al. 2009). Emerging evidence suggests that C-circles arise from replication fork collapse resolution through BIR in a RAD52-independent way (Zhang et al. 2019a; Zhang et al. 2019b). Indeed, ALT cells have been shown to undergo increased replication stress and C-circles appear following S phase suggesting that they arise from replication fork collapse resolution, which is also thought to limit genomic instability in these cells (Wang et al. 2019; Zhang et al. 2019b).

In addition, T-circles are predominantly double stranded circular telomeric DNA, thus differing from C-circles, and are commonly found in ALT cells, but are not specific to the mechanism. They are thought to arise from the t-loop excision at the distal end of the chromosome, releasing the circular t-loop resulting in a blunt ended chromosome which is highly recombinogenic and may also play a significant role in the ALT phenotype by promoting HDR at telomeres (Henson et al. 2002).

1.3.5.5 Loss of ATRX/DAXX/H3.3 complex components

The loss of ATRX is the most documented genetic mutation in relation to the ALT mechanism, however, mutations in DAXX and H3.3 have also been reported indicating that the lack of incorporation of histone H3.3 at telomeres and pericentromeric regions is important in the initiation and/or maintenance of ALT (Heaphy et al. 2011a; Bower et al. 2012; Clynes et al. 2013; Li et al. 2019). Most studies examine the loss of ATRX as a hallmark of ALT and the impact the loss of ATRX has on cells will be more extensively discussed in sections 1.4 and 1.5.

1.3.5.6 Genomic instability

The ALT mechanism is associated with a background level of instability. Indeed, ALT tumours are associated with complex karyotypes which arise from crisis however, ALT cells do not stabilise and telomeres remain dysfunctional during malignant proliferation, notably due to an increased level of replication stress (Lovejoy et al. 2012; Li et al. 2019). In addition, telomere elongation through the ALT mechanism affects other parts of the genome, notably at minisatellite MS32 present interstitially in chromosome 1 (Jeyapalan et al. 2005). This region is commonly unstable in ALT cells and the rate of instability appears to correlate with

levels of APB signal, indicating a quantifiable measure of ALT activity. It is important to note that not all minisatellites across the genome are affected and it is not clear why some regions are subject to instability as the location of the minisatellite does not appear to affect the rate of instability (Jeyapalan et al. 2005). In addition, the BLM protein appears to stabilise MS32 instability in ALT cells and plays an important role in controlling levels of instability to ensure cell viability (Mendez-Bermudez et al. 2012).

1.3.5.7 Other

Many pathways have been associated with the ALT mechanism in recent years but are not systematically affected in all ALT specimens and are therefore not specific to ALT, but contribute to the phenotype nonetheless.

1.3.5.7.1 cGAS-STING pathway

The cGAS-STING pathway is involved in the detection of extrachromosomal DNA in the cytoplasm of cells, endogenous and exogenous, which in turn initiates the expression of interferon-β thus resulting in cell death (Chen et al. 2016). Firstly, cGAS recognises the pathogenic DNA and triggers the pathway through the production of cGAMP which in turn activates the STING protein. STING is anchored to the endoplasmic reticulum, and when activated, it triggers the phosphorylation of IRF3 which migrates to the nucleus to induce interferon-β expression (Chen et al. 2016). More recently, this pathway was found to be mutated in ALT positive cells with the striking complete loss of the STING protein despite an unmutated genetic sequence (Chen et al. 2017). This in turn allows for the accumulation of C-circles in the cytoplasm of ALT cells without the initiation of this pathway. This event may therefore be required early in the process of ALT initiation and maintenance.

1.3.5.7.2 Loss of TP53

The loss of TP53 in cancer has been well documented and occurs at a high rate in cancers generally (Hollstein et al. 1991). The loss of TP53 has been suggested to be required early in the process of ALT initiation and in accordance with this hypothesis, ALT appears to occur more readily in Li-Fraumeni syndrome (LFS) tumours (Bryan et al. 1997; Henson et al. 2002). Indeed, patients with LFS, have a dominant mutation in the *TP53* gene, which encodes the TP53 protein, pre-disposing them to developing rare forms of cancers such as osteosarcoma and STSs, which often upregulate ALT for survival (Mirabello et al. 2015; Kratz et al. 2017). Loss of TP53 has also been associated with an increased rate of HDR further supporting the potential requirement for TP53 loss early in the process of ALT initiation (Bertrand et al. 1997;

Saintigny and Lopez 2002). The status of TP53 is not readily assessed in the context of ALT and this hypothesis still remains to be verified on a larger scale.

1.3.5.7.3 Increased levels of TERRA

Chromatin is less condensed in ALT cells than in normal and telomerase positive cells due to the inability to incorporate histone H3.3 upon loss of ATRX (Episkopou et al. 2014; Pickett and Reddel 2015). This in turn exposes telomeres to transcription activity, as incorporation of H3.3 at telomeres silences expression of these regions, and therefore, levels of TERRA are increased in ALT positive cells (Goldberg et al. 2010; Episkopou et al. 2014). TERRA and ATRX compete for binding sites on telomeres therefore a loss of ATRX favours TERRA localisation at telomeres (Chu et al. 2017). In addition, TERRAs have been found to accumulate at short telomeres in ALT cells forming R-loops which in turn initiate HDR-mediated repair and elongation alongside an increase in RNAseH1 protein involved in the resolution of R-loops (Arora et al. 2014; Graf et al. 2017). The FANCM complex, involved in numerous key cellular processes such as replication and DNA repair, has also been shown to resolve replication stress in ALT positive cells by disrupting R-loops, and its loss results in increased DNA damage response and repair pathways to resolve replication stress alongside a stark increase in C-circle formation due to an increase in replication fork collapse, further suggesting that C-circle formation originates from this event (Xue et al. 2015; Pan et al. 2017).

1.3.5.7.4 C-rich overhang

The presence of long G-rich overhang at the ends of telomeres is required for the formation of the t-loop; protector of the natural ends of chromosomes (Griffith et al. 1999). However, the presence of long C-rich overhangs is rare and has recently been shown to be increased in ALT positive cells and seen as a consequence of HDR-mediated telomere maintenance (Oganesian and Karlseder 2011). These overhangs appear to be generated following a sudden telomere truncation thus promoting HDR-mediated repair, perhaps through t-loop excision subsequently generating T-circles (Oganesian and Karlseder 2013). More recently, these structures have been shown to result from replication fork collapse resolution and therefore to be a consequence of ALT upregulation and an adaptation to stress as well as prevention of genomic instability in the context of ALT (Zhang et al. 2019b).

1.4 Alpha-thalassemia mental retardation syndrome X-linked (ATRX)

1.4.1 *ATRX* gene and protein structure

The *ATRX* gene is located on the q arm of chromosome X and is composed of 35 exons which span approximately 300 kb and encodes a 282 kiloDalton (kDa) protein (Picketts et al. 1996; Villard et al. 1997). The ATRX protein is a member of the Switch2/sucrose non-fermentable2 (SWI2/SNF2) family and has the distinctive feature of having an ATRX-DNMT3-DNMTL (ADD) domain, which enables DNA methyltransferase activity, alongside an C-terminal containing seven helicase domains conferring its helicase/ATPase activity, distinctive of the SNF2 family of proteins, which enables energy dependent histone-DNA contact (Picketts et al. 1998; Iwase et al. 2011; Mitson et al. 2011). In addition, it has direct binding sites for the death domain associated protein (DAXX), the heterochromatin protein 1α (HP1α), the enhancer of zeste homolog 2 (EZH2) and Methyl CpG binding domain 2 (MeCP2) protein interactions (figure 1.5) (Lewis et al. 2010; Berube 2011; Wang et al. 2017). ATRX also binds histone variants H3.3, within its ADD domain, and macroH2A (mH2A), important for recruitment, incorporation and inhibition of specific histones at repetitive heterochromatin (figure 1.5) (Ratnakumar et al. 2012).

Figure 1.5: Schematic representation of the ATRX protein.
Starting from the N-terminal to the C-terminal: in orange the ADD domain; in green the helicase domain; in red the binding sites for direct interaction with proteins; and in purple the binding domains for interactions with histones. Adapted from Cardoso et al. 1998; Nan et al. 2007; Ramamoorthy and Smith 2015; Wang et al. 2017.

1.4.2 ATRX syndrome

The ATRX syndrome was first described in 1981 suggesting a new type of X-linked mental retardation associated with α-thalassemia rather than a coincidence of the occurrence of two separate diseases within an individual (Weatherall et al. 1981). It was later confirmed that patients suffering from this disease had an intact *α-globin* gene and that the mRNA was reduced significantly, indicating that the α-thalassemia was a clinical observation of the ATRX disease (Wilkie et al. 1990). A further study isolated the *ATRX* gene as being commonly mutated in 26 cases, however, very little was known about the role of this protein in normal

cellular mechanisms (Gibbons et al. 1995). The symptoms of the ATRX syndrome are an indicator of the normal role of ATRX and are as follows: mental retardation resulting in lack of speech and mobility in most cases; characteristic facial anomalies such as small ears and inability to feed; gonadal anomalies resulting in ambiguous genitalia; and a type of α-thalassemia (Wilkie et al. 1990; Gibbons 2006). Due to the overlap in symptoms between varying X-linked syndromes, the ATRX syndrome is often mis-diagnosed, however, it is solely caused by a mutation in the *ATRX* gene, and over 100 mutations have been identified in the ADD and helicase domains resulting in varying severity of symptoms (Gibbons et al. 2008).

Due to the chromosomal location of the *ATRX* gene on the X chromosome, this syndrome mostly affects males although females can be carriers whilst exhibiting no symptoms due to skewed X chromosome inactivation (Gibbons et al. 1992). The ATRX syndrome is a very rare genetic disease and the incidence rate has not been determined although 168 cases were reported in 2006 (Gibbons 2006). There is no cure and patients are carefully monitored regularly. Due to very low numbers of patients with this syndrome and the low survival beyond 20 years of age, it is hard to address the question of predisposition to cancer although the risk would be expected to be increased, like the aforementioned genomic instability syndromes which all predispose patients to cancer. Recent studies have nonetheless published data on cases of ATRX patients developing osteosarcoma exclusively (Ji et al. 2017; Smolle et al. 2017; Masliah-Planchon et al. 2018). It is interesting to note that these case studies always reported the occurrence of osteosarcoma but not other types of sarcomas that commonly upregulate ALT for survival such as leiomyosarcoma.

1.4.3 ATRX role

ATRX has been associated with numerous cellular processes, primarily as a chromatin remodeler by incorporation of histone variant H3.3 at repetitive heterochromatin which plays a role in heterochromatin stability, DNA replication and regulation of gene expression. The roles of ATRX are described below and summarised in figure 1.6.

1.4.3.1 Genomic stability

The ADD domain of ATRX recognises and binds the H3 family of histones, notably when they are tri-methylated at lysine 9 (H3K9me3) and/or unmethylated at lysine 4 (H3K4me0), at repetitive heterochromatin regions of the genome (Dhayalan et al. 2011; Iwase et al. 2011).

Alternatively, ATRX can be recruited by HP1α, a direct interactor of ATRX, which recognises H3K9me3 and binds to it whilst it can also induce the tri-methylation process to recruit ATRX (Kourmouli et al. 2005). ATRX then recruits and directly binds the DAXX helical bundle (DHB) domain of DAXX, a H3.3 specific chaperone, and together they incorporate histone H3.3 at GC-rich repetitive regions of the DNA such as telomeres and centromeres (Drane et al. 2010; Lewis et al. 2010). ATRX and DAXX, amongst many other proteins, have been found to reside in PML nuclear bodies (PML NBs), which are bound to the nuclear matrix, and suggest a role of these PML NBs in the regulation of H3.3 incorporation (Xue et al. 2003; Tang et al. 2004a). Overall, this procures stability to these regions and condenses the heterochromatin further to prevent the transcription of telomeres into TERRAs (Chu et al. 2017). Loss of ATRX results in telomere dysfunction and an altered chromatin state, which results in a higher incidence of repair of chromosome ends through the HDR pathway alongside an increased rate of TERRA generation (Episkopou et al. 2014; Graf et al. 2017). Histone H3.3 can also be incorporated by the histone chaperone HIRA, with help from other proteins such as RPA, in other parts of the genome, including actively transcribed genes, and this mechanism is therefore unaffected by the loss of ATRX (Chow et al. 2005; Zhang et al. 2017).

Although ALT is associated with HDR-mediated repair and loss of ATRX results in a higher incidence of repair through HDR, the deposition of the H3.3 histone by the DAXX/ATRX complex also appears to play a role in SCEs. Indeed, ATRX intervention in this process occurs following RAD51 action in HDR repair and loss of it reduces levels of SCEs (Juhasz et al. 2018). Overall, SCEs occur more frequently at telomeres than in other parts of the genome, suggestive of a higher level of DNA damage at those sites, and ALT cells have shown increased rates of such events (Dunham et al. 2000; Londono-Vallejo et al. 2004; Rudd et al. 2007; Graf et al. 2017). ATRX may therefore play a role in interstitial SCEs but not telomere SCEs.

1.4.3.2 Prevention of replication stress

G quadruplexes are secondary structures that form within G-rich sequences, such as telomeres and centromeres, and present a barrier to normal cellular processes such as replication or transcription (Rizzo et al. 2009; Rhodes and Lipps 2015). ATRX plays a key role in preventing rather than resolving replication fork stalling or collapse resulting from G4s. ATRX has been found to localise to G-rich regions of the genome and is thought to incorporate histone H3.3 in complex with DAXX to resolve these secondary structures so as to remove the replication barrier (Law et al. 2010; Clynes et al. 2013; Clynes et al. 2015). Loss of ATRX results in a higher incidence of G4 structures and consequently incomplete

replication through replication fork collapse and an increase in HDR-mediated repair of replication stress (Costes and Lambert 2012; Clynes et al. 2014; Wang et al. 2019).

Another type of secondary structure which poses a barrier to DNA replication is the R-loop which is a DNA-RNA hybrid formed when TERRA binds to telomeric DNA. ATRX has been shown to migrate to actively transcribed telomeres to suppress the formation of R-loops (Nguyen et al. 2017). The exact mechanism is unclear, however, R-loops have been shown to facilitate the formation of G4 structures thus potentially recruiting ATRX to resolve these by incorporation of H3.3 (Duquette et al. 2004; Nguyen et al. 2017). In addition, ATRX also competes with TERRA to bind to DNA thus further inhibiting the formation of R-loops (Chu et al. 2017).

1.4.3.3 Regulation of gene expression

1.4.3.3.1 Negative gene expression regulation

ATRX can repress gene expression such as by interaction with the PRC2 complex and more precisely with the EZH2 protein (Cardoso et al. 1998). The PRC2 complex is involved in chromosome X inactivation, which works by tri-methylating histone H3 at lysine 27 (H3K27me3) thus resulting in repressed chromatin. ATRX recruits the complex through interaction with EZH2 and facilitates the loading of the complex onto the chromosome being silenced, although the exact interaction is not clearly understood (Sarma et al. 2014). Loss of ATRX results in mis-localisation of the PRC2 complex. In addition, the role of ATRX in sexual differentiation is unclear although it has been hypothesised that it is involved through its interaction with EZH2 and subsequently, loss of ATRX results in ambiguous genitalia observed in ATRX patients (Tang et al. 2004b).

1.4.3.3.2 Activates gene expression up stream of G-rich sequences

ATRX plays a role in regulating gene expression, such as the gene coding for the α-globin protein. Indeed, this gene is located downstream to a G-rich region and as previously mentioned, these regions are prone to forming G4 structures and the resolution of these involves the binding of ATRX and subsequently the incorporation of the histone H3.3. This in turn can drive gene expression of genes located nearby (Law et al. 2010). It is also thought that ATRX can inhibit histone m2HA deposition close to the *α-globin* gene to favour the H3.3 deposition, as these have been shown to be mutually exclusive (Ratnakumar et al. 2012). The loss of ATRX results in an increase in G4 structures leading to a reduction in α-globin expression resulting in α-thalassemia observed in ATRX patients. ATRX appears to function in

this manner in a variety of genes in mouse models such as *Nlgn4*, *Dhrsx* or *Csf2ra* all involved in brain function further supporting the impact ATRX has on gene expression and providing further explanation for the mental disabilities observed in ATRX patients (Levy et al. 2008; Levy et al. 2015).

1.4.3.4 Structural integrity of chromatin during division

Cohesion of sister chromatids occurs after S phase and is resolved prior to mitosis for appropriate segregation of chromosomes during mitosis. This mechanism facilitates SCEs for the repair of DSBs through the HDR pathway following replication (Sjogren and Nasmyth 2001). Tankyrase-1 is the main protein involved in resolving cohesion at telomeres, through its interaction with TRF1, and its action can be inhibited by direct binding of histone variant macroH2A1.1 (mH2A1.1) (Smith et al. 1998; Dynek and Smith 2004). ATRX acts by binding to mH2A1.1 preventing it from binding to tankyrase-1 (Ratnakumar et al. 2012). Telomere cohesion is often reported in ALT positive cells as the loss of ATRX means mH2A1 inhibits Tankyrase-1 action thus resulting in unresolved cohesion and inducing delayed and prolonged mitosis, although cells ultimately progress through the cell cycle eventually (Ritchie et al. 2008; Ramamoorthy and Smith 2015). In accordance with this, ALT cells also show an increase in mH2A1 deposition following replication stress to favour HDR-mediated repair (Kim et al. 2019).

1.4.3.5 Brain development

ATRX has a MeCP2 binding site within its helicase domain which enables direct interaction between the two proteins. MeCP2 is expressed in all tissues but at higher levels in the brain where it plays a major role in neuronal maturation (Kishi and Macklis 2004). It detects and binds to methylated CpG islands, an action that appears to be essential for brain development and function, as it enables the activation of transcription of specific genes, interestingly, including the *ATRX* gene (Marano et al. 2019). Together, ATRX and MeCP2 accumulate at chromocentres, which are nuclear structures containing pericentric heterochromatin (PCH), to ensure stability (Nan et al. 2007; Marano et al. 2019). These structures are instrumental in neuronal differentiation through the clustering of these chromocentres via MeCP2 and ATRX action (Bertulat et al. 2012). These structures are thought to localise within inactive regions of the genome, notably at the nuclear periphery, to ensure silencing of genes in those regions (Wijchers et al. 2015). Interestingly, loss of ATRX did not affect PCH organisation in mature neurons but rather in undifferentiated stem cells, thus indicating a key role in brain cell maturation (Marano et al. 2019). Overexpression of

ATRX also led to neuronal defects and was associated with a high level of embryonic deaths, suggesting that a tight control of ATRX is required for appropriate brain function (Berube et al. 2002).

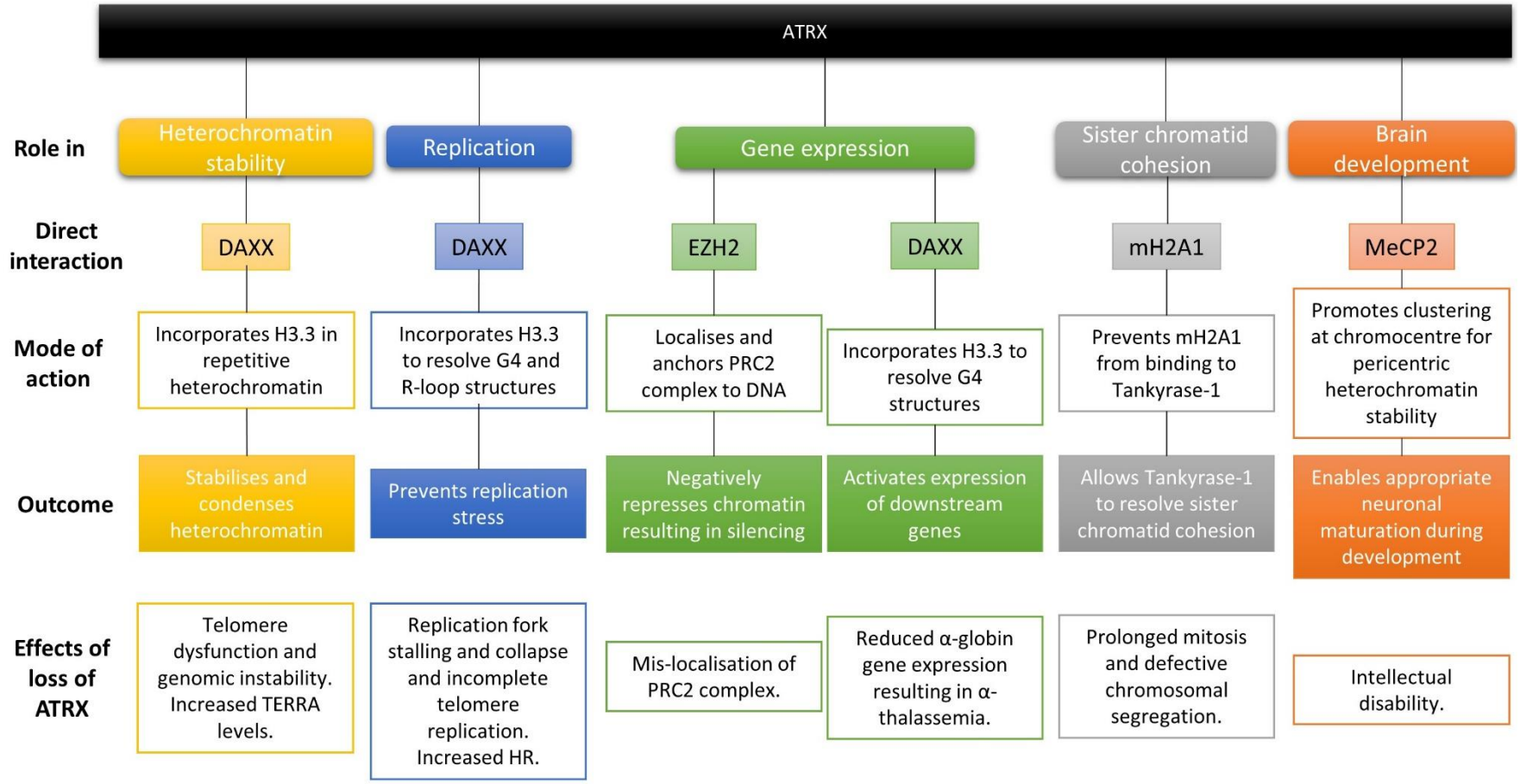


Figure 1.6: Role of ATRX.
Diagram summarising the role of ATRX in heterochromatin stability, DNA replication, gene expression, sister chromatid cohesion and brain development.

1.5 ATRX and ALT

1.5.1 Proportion of ALT cancers with an ATRX mutation

The status of ATRX has often been reported in studies evaluating the prevalence of ALT in cancers as it is often mutated (64% overall table 1.2). In contrast, the DAXX status is not always indicated as it occurs at a lower frequency, such as 4% in liposarcomas versus 78% loss of ATRX (Lee et al. 2015a). Interestingly, primary pancreatic neuroendocrine tumours (PanNETs), have an equal proportion of ATRX and DAXX mutations and appear to be mutually exclusive in the majority of cases, indicating the possibility of a more important role of DAXX in PanNET ALT upregulation and progression (Heaphy et al. 2011a; Kim et al. 2017; Singhi et al. 2017). Furthermore, the more common loss of ATRX in cancer could be due to its wider range of function and the role it plays in a DAXX-independent way, such as chromosome cohesion ensuring appropriate segregation prior to cell division (Ritchie et al. 2008).

Type	**ALT+**	**ATRX-**	**ALT+ ATRX-**	**DAXX-**	**Study**
Leiomyosarcoma	51	27/51	53%	0/51	(Liau et al. 2015b)
	33	24/33	73%	-	(Yang et al. 2015)
Total	**84**	**51/84**	**61%**		
Liposarcoma	23	18/23	78%	1/23 (4%)	(Lee et al. 2015a)
Osteosarcoma	12	7/12	58%	-	(Chen et al. 2014)
All sarcomas	119	76/119	64%		
OTHER					
Astrocytoma	53	29/53	55%	-	(Abedalthagafi et al. 2013)
Primary pancreatic neuroendocrine	98	39/98	40%	30/98	(Singhi et al. 2017)
	59	23/59	39%	28/59	(Kim et al. 2017)
	41	9/41	22%	10/41	(Heaphy et al. 2011a)
Total	**198**	**71/198**	**36%**	**68/198 (34%)**	
ALL					
	370	**176/370**	**48%**		

Table 1.2: Proportion of ATRX and DAXX loss in ALT positive cancer subtypes.

1.5.2 Role of ATRX in cancer

ATRX is more commonly mutated in ALT positive cancers than DAXX, most probably due to the role it plays across multiple cellular processes in a DAXX-independent way. ATRX appears to be lost early in the process of ALT upregulation and results in the observation of hallmarks associated with ALT, such as APB formation and C-circle generation. The loss alone of ATRX can initiate the mechanism of ALT but cannot maintain it thus implying the involvement of other processes in this pathway (Brosnan-Cashman et al. 2018). In addition, the loss of ATRX induces the ALT mechanism indirectly as within cells lacking ATRX activity, the chromatin structure at telomeres is altered due to a lack of H3.3 incorporation which in turn induces replication stress and HDR-mediated repair which drives the ALT phenotype (Episkopou et al. 2014; Graf et al. 2017). When ATRX is lost in the context of cancer, the prognosis varies according to the cancer. Indeed, the loss of ATRX is associated with a poorer prognosis in sarcomas (P value of 0.04 upon comparison of survival with or without an ATRX mutation in leiomyosarcoma) but not in astrocytic tumours (P value of 0.02 upon comparison of time until treatment failure), although the reason for this remains unknown (Wiestler et al. 2013; Yang et al. 2015; Ren et al. 2018).

1.5.3 Current limitations

1.5.3.1 Diagnostic and prognostic test

There are currently no diagnostic tests available to detect the ALT status of tumour specimens, unlike the TRAP assay to assess the presence of telomerase (Wright et al. 1995). The ability to detect hallmarks in circulating blood would be appealing as it would be less invasive than a biopsy as well as allow for the detection in cancers hard to reach such as glioblastoma or PanNets. It has been reported that C-circles are readily detected in the plasma of patients with ALT positive cancers thus providing a tool to be introduced at the diagnosis level (Henson et al. 2009). C-circles have shown to be more consistently upregulated in ALT cancers than APBs thus providing a more robust assay for determining ALT status (Henson et al. 2005; Pompili et al. 2017). In addition, with the development of novel drugs targeting telomerase, a test to detect ALT activity would be necessary to separate patients and could be used further in the future once a treatment for ALT tumours is developed. Furthermore, as previously mentioned, the upregulation of ALT combined with the loss of ATRX is associated with a poorer prognosis in STS thus enabling a further

classification of samples to provide the best line of treatment (Ren et al. 2018). This can be similarly applied to astrocytic tumours which show an inverse correlation with ALT upregulation and ATRX loss and the status of ATRX has been readily reported in patients with ALT positive astrocytic tumours (Wiestler et al. 2013; Grandin et al. 2019).

1.5.3.2 Therapeutic avenue

There are no ALT specific treatments as of yet, but multiple avenues represent an attractive target for ALT positive cancers. As described above, HDR is required for elongation of telomeres and subsequently for long-term survival. We know that both the RAD51-dependent and independent BIR pathways are used so perhaps targeting these and rendering the cancerous cells HDR-deficient would induce cell death, upon accumulation of damage associated with heterogeneous telomere lengths and replication stress, although no clinical data has been published regarding the use of DDR inhibitors in ALT cancers. In addition, G4 structures have been shown to accumulate in ALT cells that lack ATRX and stabilisation of G4 structures has shown promising results in the laboratory. Compounds targeting these structures have entered phase I of clinical trials as they prevent the resolution of G4s and therefore increase replication stress resulting in cell death (Rizzo et al. 2009; Asamitsu et al. 2019). More recently, Trabectedin has shown exciting results as it interferes with numerous cellular processes such as transcription and replication by direct interaction with DNA thus preventing the binding of other proteins (D'Incalci and Galmarini 2010). Treatment of various ALT positive cancer cell lines with Trabectedin, such as U2OS and SAOS2 which are two common ALT positive cell lines used in the literature, showed an increased sensitivity to the drug in cells with higher levels of C-circles (Pompili et al. 2017). This once more confirms that ALT activity correlates with C-circle intensity and the further need for stratification of patients according to C-circle intensity. Due to differences in response to the drug according to C-circle intensity, and potentially other hallmarks of ALT, it does suggest a varying degree of ALT activity which needs to be explored. Trabectedin has been used in various phases of clinical trials and has shown promising results in STS (Gordon et al. 2016). Finally, conflicting results regarding the sensitivity to ATR targeting have been published. ATR is involved in DSB repair by phosphorylating appropriate proteins for subsequent recruitment of repair proteins. Initially, ALT cells were shown to be highly sensitive to ATR inhibitors and resulted in cell death (Flynn et al. 2015). Another study showed no consistent increase in sensitivity but rather a cell line specific response thus making it difficult to generalise the overall reaction of ALT positive cells to this treatment method and debating the reaction of each sample to this specific drug (Deeg et al. 2016). This suggests a complex mechanism with

variability in hallmarks between cell types and even cell lines which needs to be understood further in order to develop a targeted and efficient therapeutic tool.

1.5.3.4 Understanding the upregulation and maintenance of ALT

As mentioned previously, the loss of ATRX is common but not systematic in the ALT mechanism and ATRX and DAXX mutations make up 70% of known mutations in relation to ALT upregulation (table 1.2). Despite these constituting the majority, a substantial percentage of tumour samples have an intact ATRX and DAXX whilst being ALT positive implying the involvement of other genes and pathways in the process. Therefore, the study of cellular transition from normal to malignant through the upregulation of ALT is key to gaining insight into the activation and the maintenance of this mechanism, in order to develop robust diagnostic, prognostic and potentially therapeutic tools. Furthermore, despite uncovering the hallmarks of ALT, such as the generation of C-circles and APBs, it is not yet fully understood how or why these arise and if they are required for initiation and/or maintenance of ALT. Furthermore, only putative HDR-mediated recombination events have been hypothesised and the mechanism of action of ALT still needs to be addressed fully. Understanding the sequence of events required for ALT upregulation would be significant for staging tumours as well as for understanding the main drivers of the ALT mechanism. This in turn would enable the development of novel therapies to target ALT positive tumours specifically.

1.6 Aims of the project

The main aim of this thesis was to investigate how the loss of *ATRX* combined with telomere stress, through the initiation of a telomere-driven crisis, affects different cell models in their ability to escape crisis and how this may impact genomic evolution. The individual aims were as follows:

- Assess the ability of primary human fibroblasts and telomerase positive epithelial cancer cells to escape a telomere-driven crisis in the absence of ATRX at both the single cell and polyclonal population levels
- Establish the telomere maintenance mechanism used to achieve replicative immortality in each escaping clone
- Investigate the involvement of other pathways in ALT initiation and maintenance, such as the cGAS-STING pathway
- Examine the effects the loss of ATRX has on the overall genome and if the initiation of ALT is associated with different rates of changes to the genome compared to clones upregulating telomerase
- Using Single Telomere Length Analysis (STELA) to establish the occurrence of elongation in the context of ALT upregulation and maintenance
- Address if telomeres become elongated once they reach a specific length threshold, or if it affects all telomeres irrespective of length
- Establish if telomeres are elongated to specific lengths
- Sequence ALT telomeres to understand and characterise the elongation events associated with ALT maintenance for survival as well as inferring mechanistic information

Chapter 2

Materials and methods

2.1 Materials

2.1.1 Equipment

The equipment used was sourced from various manufacturers: centrifuges (Microcentaur and centaur 2/Pico 21 and Fresco 21) from MSE/Heraeus Thermo Fisher Scientific; hot blocks from Grant, Jencons PLS and Techne; water baths from Grant; hybridisation ovens (Hybaid shake and stack/Maxi 14) from Thermo Fisher Scientific/Quanta Biotech; transilluminators from Flowgen and labtech; tissue culture cabinets (Herasafe ks) from Thermo Fisher Scientific; Incubator (Heracell 150) from Heraeus; microscope (Primovert) from Zeiss; roller mixer from Stuart Scientific; rocker (35) from Labnet; rotator (Bibby) from Stuart Scientific; shakers (R100/miniorbital shaker) from Lukcham/Stuart Scientific; aspirators (Sam12) from MGE; vortex form Thermo Fisher Scientific; power packs for gel electrophoresis from Bio-Rad; cooling pumps from KNF lab and Grant; gel electrophoresis tanks from CBS, MultiSUB Choice and Bio-Rad.

2.1.2 Plastic ware

The laboratory materials used were sourced from various manufacturers: tissue culture flasks, dishes and plates from Falcon and Greiner; cryovials (simport) from Elkay; 1 ml-25 ml stripettes from Fisher; PCR strips from Thermo Fisher Scientific and VWR; falcon tubes from Sarstedt; 0.5-2 ml tubes from StarLab, Sarstedt, Ambion and Alpha Laboratories; pipette tips from Gilson.

2.1.3 Chemicals

Commonly used chemicals were sourced from various manufacturers: PBS, NaOH, SDS (solid), NaCl (solid), Tris base, HCl, Acetic acid, EtOH, Boric acid, Methanol and Glycine from Thermo Fisher Scientific; EDTA, Tris HCl, SDS (solution), SSC, NaCl (solution), BSA (solid) and Tween 20 from Sigma.

2.2 Methods

2.2.1 Tissue culture

2.2.1.1 Cell growth and media

All cells were grown at 37 degrees Celsius (°C) in 5% CO_2.

The HCT116$^{ATRX-/-}$ cell line was provided by Professor Eric Hendrickson (University of Minnesota, Minneapolis) and cultured in Gibco McCoys 5A medium (Thermo Fisher Scientific) complemented with 10% v/v foetal calf serum (FCS) (Thermo Fisher Scientific), 1% Penicillin and Streptomycin (stock 10,000 U/ml and 10 mg/ml respectively; Sigma), 1% L-Glutamine (stock 200 mM; Sigma), 1% G418 (stock 400 µg/ml; Thermo Fisher Scientific). The DN-hTERT clones were cultured in the same medium with added 1% puromycin (stock 250 µg/ml; Calbiochem)

The HCA2 WT cells were provided by Professor James Smith (Texas medical centre, Houston) and were cultured in Gibco Dulbecco's Modified Eagle's Medium (DMEM) (Thermo Fisher Scientific) complemented with 10% v/v FCS (Thermo Fisher Scientific), 1% Penicillin and Streptomycin (stock 10,000 U/ml and 10 mg/ml respectively; Sigma), and 1% L-Glutamine (stock 200 mM; Sigma). The HCA2$^{HPV\ E6E7}$ cell line and clones were cultured in the same medium with added 1% G418 (stock 400 µg/ml; Thermo Fisher Scientific).

The U2OS cell line was provided by Professor Richard Gibbons (Weatherall Institute of Molecular Medicine, Oxford) and was cultured in Gibco DMEM (Thermo Fisher Scientific) complemented with 20% v/v FCS (Thermo Fisher Scientific), 1% Penicillin and Streptomycin (stock 10,000 U/ml and 10 mg/ml respectively; Sigma) and 1% L-Glutamine (stock 200 mM; Sigma).

2.2.1.2 Trypsinisation

Cells were passaged when flasks were 80-90% confluent, every 7 days on average. The medium was aspirated off and 1 x trypsin (Thermo Fisher Scientific) was added to wash off any remaining medium (300 µl for 24-well plates; 600 µl for 6-well plates; 1 ml for T25 flasks; and 2 ml for T75 flasks). The trypsin was aspirated, and the same amount was once more added to inhibit cell adhesion to the plastic. Cells were returned to the incubator for 5 minutes. Once cells were detached, fresh medium (1.5 ml for 24-wells; 2.4 ml for 6-wells; 4 ml for T25 flasks; and 8 ml for T75 flasks) was added to the cells to inhibit trypsin action.

2.2.1.3 Counting cells

Following trypsinisation of cells, an aliquot was taken to count the number of cells in the sample through fluorescence image cytometry on the NC-3000 (Chemometec) (50 µl for large volumes: T25 and T75 flasks; and 20 µl from small volumes: 6-well plates). No cell count was done on samples taken from a 24-well plate due to the small sample size, instead, 1 ml was passaged to a 6-well plate and 800 µl were used for DNA extraction. 2.63 µl and 1.05 µl of Solution 13 (30 µg/ml Acridine Orange, Sigma; 100 µg/ml DAPI, Sigma) was added to 50 µl and 20 µl samples respectively and 10 µl were transferred to an NC-slide A8 (Chemometec) and inserted into the cell counter. Upon cell counting, 1 x 10^5 – 4 x 10^5 cells were plated in a new flask depending on the cell type; 5 x 10^3 – 1 x 10^6 cells were used for DNA extraction; and 4 x 10^5 – 5 x 10^6 cells were stored in liquid nitrogen if possible. In addition, 5 x 10^5 – 5 x 10^6 cells were used for Western or TRAP protein extraction when needed.

2.2.1.4 Cell freezing and thawing

To freeze samples, cells (4 x 10^5 – 5 x 10^6) were centrifuged at 1,000 rpm for 5 minutes and excess medium was aspirated to a final volume of 0.5 ml. 0.5 ml of freezing mix (80% FCS; 20% DMSO, Sigma) was then added and the total volume (1 ml) was transferred to a cryovial and put in a Mr. Frosty containing isopropyl alcohol and stored at -80 °C for a minimum of 8 hours. Samples were then transferred to liquid nitrogen (-196 °C) for long-term storage.

To use frozen samples, cells were thawed in a 37 °C water bath after which 10 ml of appropriate growth medium was added drop-wise to remove the DMSO. Cells were centrifuged at 1,000 rpm for 5 minutes and excess medium was removed and cells were plated in appropriate plastic ware.

2.2.1.5 Retroviral transfection

Retroviral transfections were used to transfect HCT116$^{ATRX-/-}$ cells with the dominant negative-hTERT (DN-hTERT) cassette to abrogate telomerase activity. Recombinant retroviruses containing a pBABE-puro vector (Addgene) with a DN-hTERT encoding gene and a puromycin selection gene were grown using ΨCRIP cells, gifted by Richard Mulligan (Whitehead Institute, Cambridge). HCT116$^{ATRX-/-}$ cells (5 x 10^5) were treated with polybrene (8 µg/ml; Sigma) for 1 hour at 37 °C prior to transfection. Cells were then incubated with the DN-hTERT containing retroviruses and 48 hours after transfection, cells were plated at limited dilutions (1:10; 1:100; 1:250; 1:500; 1:1000) in order to pick single cell clones. In addition, a subset of cells were incubated with pBABE-puro vector without the DN-hTERT but

with the puromycin selection gene to assess retroviral integration effects on cells. Cells that had successfully integrated the vector were selected for using puromycin 72 hours after addition of the retrovirus to the cells (2.5 µg/ml; Calbiochem) and medium containing puromycin was subsequently used for culturing of these cells.

2.2.1.6 Nucleofection and cell sorting

Nucleofections were used to knock out *ATRX* using CRISPR/Cas9 technology in HCA2 cells. The ATRX target sequence, 5'- GTTTCTGTCGGTCGCCTCAA -3', was used as the guide RNA to target exon 9 of the *ATRX* gene and ligated into a pSpCas9(BB)-2A-GFP plasmid (Addgene plasmid #48138) using Bbs1 restriction enzyme cut sites (see figure 3.4; chapter 3). Following a cell count, 1 x 10^6 cells ($HCA2^{HPV\ E6E7}$ and HCA2 WT) were centrifuged at 700 rpm for 10 minutes. Cells were then nucleofected with 2.5 ng/µl of CRISPR plasmid using the SE nucleofector kit (V4XC-1024), as per the user's manual, and program CA-137 on the Amaxa nucleofector (Lonza). Cells were kept 10 minutes at room temperature, following which, 400 µl of warm Gibco Roswell Park Memorial Institute 1640 medium (RPMI; Thermo Fisher Scientific) were added and cells were placed at 37 °C for a further 10 minutes. Cells were then transferred to a T25 flask with fresh appropriate medium.

24 hours after nucleofection, cells were trypsinised and washed twice in warm 1 x PBS and transferred to sterile FACS tubes in order to sort cells by flow cytometry according to GFP signal. For each experiment, a negative control, consisting of cells not targeted with the CRISPR, were first used to define the selection gates to remove any cells with background GFP signal. Cells were then sorted and only strong GFP intensity cells were collected using the ARIA III FACS sorter (Becton Dickinson) and cells were collected in sterile 15 ml falcons containing fresh appropriate medium. Cells were then plated at limited dilutions (1:100; 1:250; 1:500) for single cell clone picking.

2.2.1.7 Cell cloning

Single cell clones were isolated 2-3 weeks after transfection (HCT116 cells) or cell sorting (HCA2 cells). Cloning rings were placed around single cell clones and cells were washed twice with 100 µl of 1 x trypsin and incubated 5 minutes at 37 °C. Cells were then transferred to a 24-well plate containing 1 ml of fresh warm appropriate medium. When wells were confluent (approximately 5 x 10^4 – 1 x 10^5 cells), a sample was taken for DNA extraction (800 µl) and remaining cells were passaged to a 6-well plate (1 ml). When wells were confluent (1 x 10^5 – 1 x 10^6), samples were taken for cell counts and DNA extraction and cells were passaged to a

T25 flask. HCT116$^{ATRX-/-}$ cells were subsequently kept and passaged in T25 flasks. HCA2$^{HPV\ E6E7}$ and HCA2 WT were passaged to a T75 and subsequently kept and passaged in T75 flasks.

2.2.2 DNA extraction

2.2.2.1 Phenol/chloroform

For samples with more than 10^5 cells, DNA was extracted using the phenol:chloroform method as previously described (Sambrock et al. 1989). Cells were incubated for approximately 16 hours at 45 °C with 300 µl of lysis buffer (100 mM NaCl; 10 mM Tris HCl pH 8; 5 mM EDTA pH 8; and 0.5% SDS); 3 µl 10 mg/ml RNAse (Sigma); and 3 µl 20 mg/ml Proteinase K (Sigma). All subsequent work was undertaken in a fume cabinet (Monair Plus from Astec). 300 µl of phenol:chloroform (Sigma) were then added and samples were placed on a rotator for 25 minutes. Samples were centrifuged at 13,000 rpm for 5 minutes and the aqueous phase was transferred to a new tube containing 300 µl of phenol:chloroform. Samples were rotated for a further 20 minutes and centrifuged at 13,000 rpm for 5 minutes. The aqueous phase was transferred to a fresh tube and the DNA was ethanol precipitated by adding 30 µl of 3 M sodium acetate pH 5.2 (Sigma) and 900 µl of absolute ethanol (96-100%). Samples were placed at -20 °C overnight and centrifuged at 13,000 rpm for 5 minutes and excess ethanol was removed. Pellets were then washed in 70% ethanol and stored on ice for 10 minutes. Samples were once more centrifuged at 13,000 rpm for 5 minutes, excess ethanol was removed and DNA was left to air dry for approximately one hour. All samples were re-suspended in 50 µl 10 mM Tris HCl pH 8 and placed at -20 °C for long-term storage.

2.2.2.2 QIAamp DNA Micro Kit (Qiagen)

For samples with less than 10^5 cells, DNA was extracted using the Qiagen kit as per the user's manual: Protocol for isolation of genomic DNA from small volumes of blood. 1 µl of carrier RNA was added to each sample. For samples with 5 x 10^3 cells and below, DNA was eluted in 35 µl of 10 mM Tris HCl pH 8. For samples with 10^4 to 10^5 cells, DNA was eluted in 50 µl of 10 mM Tris HCl pH 8.

2.2.2.3 DNA quantification

Following extraction, samples were incubated at 37 °C on a hot block for 30-60 minutes to ensure complete solubilisation. DNA was then quantified on a QuantiFluor-ST fluorometer (Promega) using Hoechst 33258 dye. The fluorometer was calibrated using 500 ng (10 µl) of

calf thymus DNA (100 µg/ml; Bio-Rad) with 2 ml of 1 x TEN buffer (ddH_2O; 10 x TEN stock buffer: 100 mM Tris, 10 mM EDTA and 2 M NaCl, pH 7.4; and 0.1 µl/ml Hoechst dye (Bio-Rad)). DNA samples were quantified by adding 2 ml of 1 x TEN buffer with Hoechst to 2 µl of DNA. Each sample was run in triplicate from which an average was calculated.

2.2.3 Protein extraction

2.2.3.1 Extraction for Western

Cell pellets containing 5 x 10^5 to 5 x 10^6 cells were washed twice in 1 x PBS for protein extraction for Western blot analysis. Pellets were incubated for 5 minutes on ice with 2.5 x their cell volume in lysis buffer (150 nM NaCl; 50 mM Tris; 5 mM EDTA; 1% NP40 dissolved in water pH 8.0 (Fluka)); 30 µl/ml PMSF (100 mM; Sigma); 1:100 Protease Inhibitor Cocktail III (Calbiochem); 1:100 Phosphatase Inhibitor Cocktail II (Calbiochem)). Samples were then centrifuged at 20,000 g at 4 °C for 30 minutes. The supernatant containing the protein was aliquoted into fresh tubes and stored at -80 °C until required and 5 µl were diluted in 45 µl ddH_2O (1:10) for protein quantification.

2.2.3.2 Extraction for telomerase quantification

Cell pellets containing 5 x 10^5 to 5 x 10^6 cells were washed twice in 1 x PBS for protein extraction for telomerase quantification (TRAP assay). Cell pellets were re-suspended in 200 µl CHAPS lysis buffer (Millipore) and stored on ice for 30 minutes. Samples were then centrifuged at 12,000 g at 4 °C for 20 minutes. The supernatant was aliquoted and stored at -80 °C until required and 40 µl were used for protein quantification.

2.2.3.3 Protein quantification

Following extraction (for both TRAP and Western), protein was quantified on the Cytation 3 imaging reader (Biotek) using a 96-well plate. All samples were run in duplicates and 10 µl of each sample were incubated with 300 µl Coomassie Plus (Thermo Fisher Scientific) for 10 minutes at room temperature. The standards used to obtain the standard curve were varying dilutions of BSA (2 mg/ml stock, Thermo Fisher Scientific). For TRAP protein quantification the following dilutions were used: 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2 (all mg/ml); for Western protein quantification the following dilutions were used: 0.1, 0.2, 0.3, 0.4 (all mg/ml). The absorbance at the wavelength of 595 nm was measured and sample concentrations were calculated using the standard curve on Excel.

2.2.4 PCR

2.2.4.1 Oligonucleotides

All primers listed below were synthesised by Eurofins Genomics and are in the 5′ to 3′ orientation. STELA and fusion primers were designed by examining the sub-telomere sequence of various chromosome ends to ensure that primers were 20-24 nucleotides long with appropriate GC content and a minimum of a 1 bp difference with other sub-telomeres to ensure single telomere amplification. These were designed by Prof. Duncan Baird and Dr Kevin Norris (Baird et al. 2003; Britt-Compton et al. 2006; Norris et al. 2019). Screening and sequencing primers were designed by Dr Rhiannon Robinson. TRAP (method 2) primers were taken from the following reference: Mender and Shay 2015.

Application	Name	Sequence	Annealing temperature	DNA quantity per reaction
STELA	2p2	GAGCTGCGTTTTGCTGAGCAC	65°	-
	5p5	GGAGCAGCATTCTCTTCACCACAG	59°	1.25 ng
	7qK1	GGGCACTGCCTCGCTTTGA	65°	1.25 ng
	8q2	CCCTGGAAAGGACATAAATTCG	65°	2.5 ng
	9p2	CACATTCCTCATGTGCTTACG	61°	2.5 ng
	11q13B	CAGACCTTGGAGGCACGGCCTTCG	66°	-
	12qK1	TCAGCACAGACGCGGGCGGTT	65°	-
	16prev1	CACTTATTAGTTCCAGTCTCTG	56°	-
	17pseq1rev	GAATCCACGGATTGCTTTGTGTAC	59°	250 pg
	17p6	GGCTGAACTATAGCCTCTGC	59°	250 pg
	18qrev4M	CACAGGGATGGTTAGGTATCTC	59°	250 pg
	XpYpC	CAGGGACCGGGACAAATAGAC	65°	250 pg
	XpYpE2	TTGTCTCAGGGTCCTAGTG	65°	250 pg
	Teltail	TGCTCCGTGCATCTGGCATC	59-65°	
	Telorette 2	TGCTCCGTGCATCTGGCATCTAACCCT	59-65°	
Fusion	17p6	GGCTGAACTATAGCCTCTGC	62°	25 ng
	21q1	CTTGGTGTCGAGAGAGGTAG	62°	25 ng
	XpYpM	ACCAGGTTTTCCAGTGTGTT	62°	25 ng
Screening	Ax9ScrnF	AGTGGAACTGAACAAGAAGTGG	63.6°	20 ng
	Ax9ScrnR	GAAGGCACAGTTGATAAAGACACG	63.6°	20 ng
Sequencing	Ax9SeqF	CCTGTTTCCCTTTCTAATTCCC	N/A	
TRAP (method 2)	Cy5-TS	ATTCCGTCGAGCAGAGTT	N/A	
	ACX	GCGCGGCTTACCCTTACCCTTACCCTAACC	N/A	
	NT	ATCGCTTCTCGGCCTTTT	N/A	

Table 2.1: Oligonucleotides used according to application.
'-' signifies that despite optimisation and an increase in DNA concentration, no adequate amplification was obtained.

2.2.4.2 STELA

Following DNA extraction, a working stock of DNA diluted to 10 ng/µl was made for subsequent use. DNA was diluted to varying dilutions in Tris-HCl (10 mM pH 8) and 250 nM of telorette 2 (specified in table 2.1) according to the telomere specific primer used. For each sample, 4-6 reactions of 10 µl each were amplified. A PCR master mix was added to the DNA and was composed of: Taq 1 x buffer (10 x stock: 75 mM Tris HCl pH 8.8; 20 mM $(NH_4)SO_4$; and 0.01% Tween20; Thermo Fisher Scientific); 2 mM MgCl2 (Thermo Fisher Scientific); 1.2 mM dNTPs (Promega); 0.5 µM teltail; 0.5 µM telomere specific primer; 1 U Taq/PWO (10:1) (Thermo Fisher Scientific/Roche); and ddH_2O to make up 10 µl per reaction. The reactions were cycled using a Tetrad thermal cycler (Bio-Rad) at the following conditions: 94 °C for 20 seconds; 59 °C – 65 °C for 30 seconds (temperature dependent on telomere specific primer); and 68 °C for 8 minutes for 22 cycles.

2.2.4.3 Fusion

Following DNA extraction, a working stock of 25 ng/µl DNA was made for subsequent use. For each sample, 5-6 reactions of 10 µl each were amplified. For each reaction, 25 ng of DNA were incubated with Taq 1 x buffer (10 x stock: 75 mM Tris HCl pH 8.8; 20 mM $(NH_4)SO_4$; and 0.01% Tween20; Thermo Fisher Scientific); 1.2 mM $MgCl_2$ (Thermo Fisher Scientific); 1.2 mM dNTPs (Promega); 0.5 µM telomere specific primers (XpYpM; 17p6; and 21q1); 1 U Taq/PWO (10:1) (Thermo Fisher Scientific/Roche); and ddH_2O to make up 10 µl per reaction. The reactions were cycled using a Tetrad thermal cycler (Bio-Rad) at the following conditions: 94 °C for 20 seconds; 62 °C for 30 seconds; and 68 °C for 8 minutes and repeated for 25 cycles.

2.2.4.4 Screening

To assess if *ATRX* was successfully knocked out in CRISPR/Cas9 targeted HCA2 cells, a screening PCR was carried out. Following DNA extraction, 20 ng of DNA were incubated with Taq 1 x buffer (10 x stock: 75 mM Tris HCl pH 8.8; 20 mM $(NH_4)SO_4$; and 0.01% Tween20; Thermo Fisher Scientific); 2 mM $MgCl_2$ (Thermo Fisher Scientific); 1.2 mM dNTPs (Promega); 0.5 µM Ax9ScrnF primer; 0.5 µM Ax9ScrnR primer; 1 U Taq/PWO (10:1) (Thermo Fisher Scientific/Roche); and ddH_2O to make up 15 µl per reaction. The reactions were cycled using a Tetrad thermal cycler (Bio-Rad) at the following conditions: 94 °C for 20 seconds; 63.6 °C for 30 seconds; and 72 °C for 1 minute for 32 cycles.

Following PCR amplification, PCR amplicons were cleaned-up using the Monarch PCR & DNA clean-up kit (NEB) according to the user's manual and samples were eluted in 25 µl of 10 mM

Tris HCl pH 8. 8.6 µl of amplified DNA were then incubated with x 10 CutSmart buffer (NEB) and 4 U SmlI restriction enzyme (NEB) at 55 °C for 1 hour using a Simpliamp thermal cycler (Life technologies).

2.2.4.5 C-circle

Following DNA extraction, 20 ng DNA were incubated with 0.2 mg/ml BSA (NEB); 0.1% Tween 20 (Thermo Fisher Scientific); 1 mM dATP, dGTP and dTTP (Promega); 1 x ϕ29 buffer (50 mM Tris HCl; 10 mM $MgCl_2$; 10 mM $(NH_4)SO_4$; and 4 mM DTT; NEB); 7.5 U ϕ29 polymerase (NEB); and nuclease free water to make up 20 µl reactions. Reactions were incubated for 8 hours at 30 °C followed by 20 minutes at 65 °C using a Simpliamp thermal cycler (Life technologies) as previously described (Henson et al. 2009). A minus polymerase control was included for each sample using the same PCR programme and master mix without ϕ29 polymerase.

2.2.4.6 TRAP

Following protein extraction, a working stock of 30 µl at 100 ng/µl was prepared for each sample to be amplified by PCR.

2.2.4.6.1 TRAPeze XL telomerase detection kit (method 1)

200 ng of protein were incubated with 5 x TRAPeze XL reaction mix (Millipore); 1 x Titanium Taq polymerase (Clontech); and ddH_2O to make up 50 µl reactions. In addition, for all samples, 10 µl were incubated at 100 °C for 15 minutes to inactivate telomerase and similarly, 200 ng of protein were incubated with the same master mix. The controls were as follows: positive control provided by the kit at a dilution of 1:10; a minus telomerase control (2 µl of CHAPS buffer instead of sample); and a minus Taq polymerase control. The standards used to create the standard curve were varying dilutions (0.2; 0.04; 0.008; and 0.0016 amoles/µl) of TSR8 provided with the kit. Reactions were cycled using a Tetrad thermal cycler (Bio-Rad) at the following conditions: 30 °C for 30 minutes for 1 cycle; 94 °C for 30 seconds; 59 °C for 30 seconds; 72 °C for 1 minute for 36 cycles followed by 72 °C for 3 minutes and 55 °C for 25 minutes for 1 cycle.

2.2.4.6.2 TRAP assay by gel electrophoresis (method 2)

500 ng of protein were incubated with 1 x TRAP buffer (2 x stock: 40 mM Tris HCl pH 8; 3 mM $MgCl_2$; 126 mM KCl; 0.01% Tween 20; 2 mM EGTA; 0.2 mg/ml BSA; 0.1 mM dNTPs); 0.36 µM TS primer; 1 µl primer mix (stock 0.10 µM ACX primer; 0.19 µM NT primer; and 0.0025 pM TSNT internal control 5'-AATCCGTCGAGCAGAGTTAAAAGGCCGAGAAGCGAT-3'); 0.4 x

Titanium Taq polymerase (Clontech); and ddH_2O to make up 50 µl. Reactions were cycled using a Tetrad thermal cycler (Bio-Rad) at the following conditions: 25 °C for 40 minutes; and 95 °C for 5 minutes for 1 cycle; 95 °C for 30 seconds; 52 °C for 30 seconds; and 72 °C for 45 seconds for 29 cycles followed by 72 °C for 10 minutes for 1 cycle.

2.2.5 Gel electrophoresis

2.2.5.1 STELA and fusion

2 µl of 6 x Ficoll gel loading solution (5% bromophenol blue; 5% xylene; and 15% Ficoll) were added to each reaction. For both STELA and fusions, DNA amplicons were resolved on a 40 cm 0.5% low density agarose (Roche) gel made with 1 x Tris-acetate EDTA (TAE) (3 L of dH_2O and 20 ml of 60 x TAE stock: 40 mM Tris base; 20 mM acetic acid; 1 mM EDTA) and ethidium bromide (1 µg/ml; Thermo Fisher Scientific).

Up to 40 STELA reactions could be loaded at one time and up to 120 fusion reactions could be loaded per gel. Gels were run in 1 x TAE at 4 °C for approximately 16 hours at 110 V for STELA and 40 V for fusions. Alternatively, fusions were also run for 3 hours at 200 V.

2.2.5.2 Screening

2 µl of 6 x Ficoll gel loading solution (5% bromophenol blue; 5% xylene; and 15% Ficoll) were added to each reaction. Purified PCR amplicons (Sml1 treated and untreated) were resolved on a 0.7% low density agarose (Thermo Fisher Scientific) gel made with 1 x TAE and 1 µg/ml ethidium bromide (Thermo Fisher Scientific). Up to 20 reactions could be loaded at one time. Gels were run in 1 x TAE for 1-2 hours at 110 V.

2.2.5.3 Western

5 µl of 3 x loading buffer (0.15 M Tris-HCl; 31% glycerol; 3.2% SDS; 3% 2-mercaptoethanol; and 0.06% bromophenol blue) were added to 5-10 µg of protein to make a total of 15 µl per sample. Samples were incubated at 100 °C for 5 minutes after which they were loaded onto Mini-PROTEAN III TGX precast gels (7.5%; 10%; or gradient 4-15% gels) (Bio-Rad). Gels were run in 1 x running buffer (400 ml dH_2O and 100 ml of 5 x running buffer stock: 0.12 M Tris; 1.25 M glycine; 17 mM SDS; at pH 8.3) for 1-2 hours at 100 V.

2.2.5.4 TRAP (method 2)

10 µl of 6 x Ficoll gel loading solution (5% bromophenol blue; 5% xylene; and 15% Ficoll) were added to each reaction. TRAP PCR products were resolved on acrylamide gels (12.5% acrylamide 19:1; 0.06% APS; 0.125% TEMED; and 0.6 x Tris-borate-EDTA (TBE)). Gels were run in 1 x TBE (400 ml of dH_2O and 100 ml of 5 x TBE stock: 0.45 M Tris base; 0.45 M boric acid; and 10 mM EDTA pH 8) for 1-2 hours at 100 V.

2.2.6 Blotting

2.2.6.1 Southern

STELA and fusion PCR products were detected by Southern Hybridisation. The bands were visualised using a transilluminator to cut the gel according to the DNA molecular weight ladders. Gels were then washed 2 x 6 minutes in depurination buffer (0.25 M HCl) followed by a 15 minutes wash in denaturation buffer (1.5 M NaCl; 0.5 M NaOH). DNA amplicons were transferred to a positively charged membrane (Hybond-XL GE healthcare) under basic conditions for 4-6 hours. The membrane was then pre-hybridised in Church buffer (0.5 M sodium phosphate dibasic, 7% SDS, 1% BSA, 1 mM EDTA) for 30-60 minutes at 56 °C.

2.2.6.2 Slot

The C-circle PCR reactions were blotted using the Bio-Dot SF apparatus (Bio-Rad) which functions by causing a vacuum to aspirate the samples through wells and allowing the DNA to bind to the positively charged membrane (Hybond-XL GE healthcare). Samples were first diluted to 150 µl with nuclease free water, 0.4 M NaOH and 10 mM EDTA and incubated at 100 °C on a hot block for 5 minutes to denature the DNA. Samples were then added to individual wells and the vacuum was applied. 500 µl of 0.4 M NaOH were then added to each well to create basic conditions for appropriate transfer of DNA to the membrane. Upon application of the vacuum, the manifold was dismantled, and the membrane was soaked in 2 x SSC and pre-hybridised in Church buffer for 1 hour at 56 °C.

2.2.6.3 Western

Protein was transferred to a PVDF membrane (Immobolin-P Millipore) following gel electrophoresis. The pre-cut PVDF membrane was firstly immersed in methanol for 3 seconds followed by ddH_2O for 2 minutes. Protein was then transferred onto the membrane in

transfer buffer (25 mM Tris; 0.19 M glycine; 20% methanol; 0.01% SDS; at pH 8.1-8.4) for 4-6 hours at 100 V at 4 °C or alternatively overnight at 30 V.

2.2.7 Hybridisation

2.2.7.1 Radiolabelling DNA probes

Radiolabelled DNA probes were synthesised with different kits and according to the user's manual. Firstly, 45 µl of TE buffer and 1 µl of 25 ng/µl DNA (telomere DNA for STELA and C-circle assay and chromosome specific DNA for fusions) were incubated at 96 °C for 5 minutes to denature the DNA followed by 5 minutes on ice. The content was then added to an Amersham Rediprime II Random Prime Labelling Bead (GE Healthcare) to which 4 µl of [^{33}P]dCTP (3,000 Ci/mmol; PerkinElmer) were added and incubated at 37 °C for a minimum of 20 minutes. 50 µl of ddH_2O and 1 µl of radiolabelled ladder (1 kb and 2.5 kb DNA) were finally added prior to use.

Alternatively, the megaprime labelling kit (Amersham GE healthcare) was used for which 43 µl of TE buffer, 1 µl of 25 ng/µl DNA (telomere DNA for STELA and C-circle assay and chromosome specific DNA for fusions), 1 µl of ladder DNA (1 kb and 2.5 kb DNA) and 5µl of primer (provided with the kit) were incubated at 96 °C for 5 minutes to denature the DNA followed by 5 minutes on ice. 10 µl of labelling buffer (provided with the kit), 2 µl of Klenow enzyme and 2-4µl of [^{32}P]dCTP (800 Ci/mmol; PerkinElmer) were added and incubated at 37 °C for a minimum of 1 hour. 36 µl of ddH_2O were finally added prior to use.

2.2.7.2 Hybridisation

The radiolabelled probe was denatured at 96 °C for 5 minutes after which 25 µl were added to each bottle containing one blot. Blots were hybridised overnight at 56 °C in a rotating hybridisation oven or alternatively for 4-6 hours at 56 °C.

2.2.7.3 Washes

Following hybridisation, blots were washed with 0.1% SDS and 0.1 x SSC wash buffer for 3 short washes followed by 2 x 15 minutes washes. STELA and C-circle blots were then dried at 56 °C and fusion blots were wrapped in cling film for optimal subsequent re-probing.

2.2.7.4 Stripping

Fusion and STELA blots were incubated with 0.2% SDS at boiling temperature for a minimum of 2 hours to strip radiolabelled probe previously applied to the blots. Blots were exposed to a phosphor screen (Amersham) and scanned on a Typhoon FLA 9500 imager (GE healthcare) to monitor for any residual signal. Stripping step was repeated until no signal was detected on the blots after which these could be re-probed with another radiolabelled probe.

2.2.8 Visualisation

2.2.8.1 STELA, fusions and C-circles

Following hybridisation with appropriate probe, STELA, fusion and C-circle blots were exposed to a phosphor screen (Amersham) for 10-24 hours and imaged using a Typhoon FLA 9500 imager (GE Healthcare) using the Phosphor filter (635 nm laser wavelength; 200 micron pixel size). Fusion blots were then stripped and re-probed sequentially to assess XpYpM:17p6:21q1 fusion events.

2.2.8.2 UV

Following gel electrophoresis of standard DNA, gels (used for ATRX screening) were visualised on a Labnet UV transilluminator for which the exposure time could be increased and photos could be saved for analysis.

2.2.8.3 Western blots

Following protein transfer, membranes were rinsed in wash buffer (1% Tween 20; and 1 x PBS). Membranes were then incubated with blocking buffer (STING antibody blocking buffer: 25 mM Tris pH 7.4; 0.15 M NaCl; 0.1% Tween 20; and 5% milk. Blocking buffer for ATRX and Vinculin: 0.2% Tween 20; 5% milk; and 1 x PBS) for 1 hour at room temperature on a rocker. Following blocking, membranes were incubated with primary antibody solution (detailed in table 2.2) in a 15 ml falcon on a roller at 4 °C overnight, or alternatively at room temperature for 1-2 hours. Membranes were then washed 5 x 5 minutes in wash buffer (1% Tween 20 and 1 x PBS) and incubated with secondary antibody solution (detailed in table 2.2) in a 15 ml falcon on a roller at room temperature for 1 hour. Membranes were finally washed 5 x 5min in wash buffer and protein was visualised with the ECL Plus kit (Thermo Fisher Scientific) for which 50 µl of solution A was added to 2 ml of solution B and poured over the membrane

and incubated for 5 minutes at room temperature. Membranes were then wrapped in Saran wrap and exposed to chemiluminescence film (Amersham) for varying amounts of time. The films were then developed on an SRX-101A Medical Film processor (Konica Minolta Medical and Graphic Inc.) in a dark room. The molecular weights of proteins were determined with the Precision Plus Protein Standards (Bio-Rad) run on each gel alongside the samples of interest.

Primary antibody concentration	Secondary antibody concentration	Antibody solution
Monoclonal mouse anti-STING (1:2,500) R&D systems Biotechne	Anti-mouse HRP (1:2,000)	25 mM Tris pH 7.4; 0.15 M NaCl; 0.1% Tween 20; and 5% milk
Polyclonal rabbit anti-ATRX (1:1,000) Santa Cruz Biotechnology	Anti-rabbit HRP (1:1,000)	1% BSA; 0.2% Tween 20; 5% milk; and 1 x PBS
Monoclonal mouse anti-Vinculin (1:10,000) Santa Cruz Biotechnology	Anti-mouse HRP (1:10,000)	1% BSA; 0.2% Tween 20; 5% milk; and 1 x PBS

Table 2.2: Summary of the primary and secondary antibody concentrations and solutions used for immunoprobing of Western blot membranes.

To strip any signal for re-probing, membranes were washed in ddH_2O for 5 minutes, 0.2 M NaOH for 5 minutes and ddH_2O for 5 minutes. Membranes could then be re-probed with another antibody following the method detailed above.

2.2.8.4 TRAP (method 1)

Following a PCR, 150 µl of Trapeze buffer (10 mM Tris-HCl pH 7.4; 0.15 M NaCl; 2 mM $MgCl_2$) were added to each PCR reaction in a 96-well plate. The excitation and emission state wavelengths for Fluorescein (485 nm and 535 nm respectively) and Sulforhodamine (585 nm and 620 nm respectively) were measured on a Cytation3 imaging reader (BioTek). All subsequent calculations were done in Excel to calculate the total product generated, indicator of telomerase activity, according to the user's manual.

2.2.8.5 TRAP (method 2)

Following gel electrophoresis, gels were incubated in 1:10,000 SYBR-gold for 10 minutes on a rocker at room temperature. Gels were then scanned using the typhoon FLA 9500 scanner using the SYBR-Gold filter (473 nm laser wavelength; 200 micron pixel size).

2.2.9 Sequencing

2.2.9.1 PacBio

To generate PCR amplicons to be sequenced using the single-molecular real-time (SMRT) PacBio sequencing platform, specific chromosome ends were amplified using STELA. A minimum of 500 ng of DNA were required for PacBio sequencing therefore, 1,600 reactions were generated for each sample. The starting concentration of all samples used, alongside the chromosome ends that were amplified are detailed in table 2.3. The master mix used was the same as described in 2.2.4.2 apart from all primers used for each sample were added to the same master mix adjusting the volume of ddH_2O accordingly to limit the amount of input genomic DNA for optimal sequencing. The reactions were cycled using a Tetrad thermal cycler (Bio-Rad) at the following conditions: 94 °C for 20 seconds; 63 °C for 30 seconds; and 68 °C for 8 minutes for 24 cycles.

Sample (Characteristic)	DNA concentration per reaction	Chromosome specific primers
$HCT116^{ATRX-/-}$ parental (control)	250 pg	7qK1; 17pseq1rev; XpYpC
$HCT116^{ATRX-/-\ DN-hTERT}$ clone 2 (following ALT-like elongation)	250 pg	7qK1; 17pseq1rev; XpYpC
$HCA2^{HPV\ E6E7}$ parental (control)	375 pg	17p6; XpYpE2
$HCA2^{HPV\ E6E7\ ATRX-/-}$ clone 21 (following ALT upregulation)	375 pg	17p6; XpYpE2
U2OS (ALT-positive control)	625 pg	17pseq1rev; XpYpE2

Table 2.3: Summary of samples sent for PacBio sequencing and the DNA concentration used alongside the amplified telomeres.

Following PCR amplification, sample reactions were pooled together and concentrated down using an ISS110 Speedvac system (Thermo Fisher Scientific) to evaporate excess water using a vacuum. All samples were then purified using AMPure XP beads (Beckman Coulter) according to the user's manual. Firstly, 1.8 x the sample volume of beads was added to each sample and mixed thoroughly. After 10 minutes at room temperature, the tubes were placed on a magnet to separate the beads from the liquid. Excess liquid was removed and the beads were washed 2 x 90 seconds in 70% ethanol. The beads were left to air-dry 10-15 minutes and re-suspended in 50 µl nuclease free water. After 5-10 minutes, tubes were placed on the magnet once more and the remaining liquid containing the cleaned DNA was retrieved.

Samples were then processed for PacBio library preparations by Mrs Joanne Morgan (NGS coordinator; Cardiff University).

2.2.9.2 Whole genome sequencing

27 DNA samples (combination of ALT positive and telomerase positive clones) were sent for whole genome sequencing using the BGISEQ-500 platform, providing paired end (2 x 100 bps) sequencing with a 15 x coverage. A minimum of 20 µl at a concentration of 1 µg per sample were sent. The QC, library preparation, sequencing and data filtering were carried out by BGI. Sequence mapping and analysis were carried out by Dr. Kez Cleal.

2.2.9.3 Screening ATRX clones

Following gel purification of CRISPR/Cas9 targeted HCA2 clones, mutant clones were sequence verified through Sanger sequencing (Eurofins genomics). 2 µl of Ax9seqF primer (10 µM; table 2.1) were added to 15 µl at a minimum concentration of 10 ng/µl per samples of interest and sent to Eurofins Genomics for sequencing. Sequenced reads were then verified for mutations at the CRISPR target site using BLAST.

2.2.10 Analysis

2.2.10.1 STELA quantification

The molecular weight of DNA fragments generated with STELA were quantified on ImageQuantTL (GE Healthcare) using the 1D gel analysis tab. Molecular weight was determined by comparison to the DNA ladder run on each gel and the length in base pairs was exported to Excel upon which the sub-telomeric DNA length between the 5' end of the sub-telomeric PCR primer and the start of the telomere repeat array was subtracted from each telomeric molecule to obtain the length of the double-stranded telomere repeat containing DNA. The mean telomere length for each sample was then calculated. Statistical analyses were carried out in GraphPad Prism 5.

2.2.10.2 Fusion quantification

The number of unique bands were counted following radioactive labelling of specific probes. When an inter-allelic event was observed, the fusion event was counted once and attributed

to the first probe that detected the event, i.e.: a fusion between 17p and 21q would be counted as a 17p fusion event as sequentially the 17p probe is used before the 21q probe.

2.2.10.3 C-circle quantification

C-circle bands were quantified on ImageQuantTL (GE Healthcare) using the array analysis tab. The intensity of the signal of each band was measured and the background (minus polymerase) signal was subtracted from each sample. All samples were then normalised to the parental cell line. Where replicates were available, error bars using the standard deviation were used to show technical variation. All calculations were done in Excel.

2.2.10.4 Western blot quantification

Western blot protein expression was quantified on ImageQuantTL (GE Healthcare) using the array analysis tab. The intensity of the signal of each band was measured and a ratio of protein of interest (STING or ATRX) to Vinculin expression was calculated for each sample. This value was then normalised to the parental value to assess the fold change in protein expression. Where replicates were available, error bars using the standard deviation were used to show technical variation. All calculations were done in Excel.

2.2.10.5 Statistical tests

T-tests, Mann-Whitney U tests, ANOVA tests, Kruskal-Wallis tests and Fisher's exact tests were all carried out in GraphPad Prism 5. The Fligner test for homogeneity of variances and common language effect size (CLES) tests were carried out in Python, the method was scripted by Dr Kez Cleal and the values were inputted by myself alongside the running of the tests.

2.2.10.6 Figure generation

All figures were generated with Adobe Illustrator and all graphs were generated with GraphPad Prism 5. Photos and schematic representations were inserted or created on PowerPoint. All figures were inserted into the text as JPEG files.

Chapter 3

Investigating the role of ATRX in primary human fibroblasts and the escape from crisis

3.1 Abstract

Cells of mesenchymal origin preferentially upregulate the ALT mechanism for survival in cancer progression. Here, the role that the loss of ATRX plays in the initiation of ALT was analysed in primary human fibroblasts.

Following the abrogation of TP53 and Rb function upon transfection with HPV E6E7 viral oncoproteins to initiate a telomere-driven crisis, HCA2 foreskin fibroblasts were nucleofected with a CRISPR vector to knockout ATRX. A total of 9 clones were successfully targeted and monitored for an escape from crisis. Of these clones, 3 clones failed to escape crisis and 6 clones immortalised by upregulating ALT as seen with a strong C-circle intensity, lack of telomerase activity and increased telomere length heterogeneity. The 6 clones that escaped crisis also displayed a complete loss of the ATRX protein, in contrast to the 3 clones that failed to escape crisis that exhibited residual ATRX protein expression despite having a mutated genetic sequence. To further explore the effects of the loss of ATRX, WT fibroblasts with functional TP53 and Rb cell cycle checkpoint proteins, were similarly targeted with the CRISPR vector to establish if the loss of ATRX alone is sufficient to initiate crisis and subsequently ALT. This experiment failed to generate ATRX knock out clones although, the mixed population showed signs of crisis through the presence of telomere fusions and precipitate, only seen during crisis.

In conclusion, the complete loss of ATRX, combined with telomere stress during crisis, via the abrogation of TP53, was sufficient to initiate ALT and this was maintained in the long-term to immortalise primary fibroblasts.

3.2 Introduction

Telomeres erode as a function of cell division and short telomeres elicit a TP53-mediated cell cycle arrest which triggers replicative senescence (Olovnikov 1973; Shay et al. 1991; von Zglinicki et al. 2005). Mutations can render the cell cycle checkpoint proteins dysfunctional, allowing cells to continue dividing despite their critically short telomeres no longer being able to protect the natural ends of chromosomes. These cells then enter a phase termed crisis during which telomere fusions occur that drive extensive genomic rearrangements resulting in instability and cell death (Counter et al. 1992; Capper et al. 2007). During crisis, the majority of cells will undergo apoptosis as a result of lethal mutations however, in some rare cases, cells will escape and gain replicative immortality leading to clonal malignant growth (Cong et al. 2002; von Morgen and Maciejowski 2018). Cancer cells become immortal by maintaining their telomere lengths by using one of two mechanisms: 85-90% of cancers upregulate the enzyme telomerase, whilst the remaining 10-15%, activate the ALT pathway (Heaphy et al. 2011b).

Little is known and understood about the initiation and maintenance of the ALT mechanism despite it being activated in 56% of sarcomas (summarised in table 1.2) (Heaphy et al. 2011b). Sarcomas are a rare form of malignancy affecting the connective tissue and account for the majority of ALT positive tumours recorded (Heaphy et al. 2011b). Most ALT positive cancers arise in cells of mesenchymal origin. MSCs differentiate to become connective tissue such as adipocytes in adipose tissue, osteocytes in the bone and chondrocytes in the cartilage (Dominici et al. 2006). MSCs require little to no telomerase activity during development or differentiation in adult life thus suggesting a pre-disposition for ALT upregulation in these cells (Parsch et al. 2004). A common hallmark observed in ALT positive cells is the loss of the ATRX protein involved in numerous cellular processes with its primary role being the incorporation of histone H3.3, alongside DAXX, in repetitive heterochromatin such as telomeres and centromeres to stabilise these structures as well as resolve secondary DNA structures (Lewis et al. 2010; Heaphy et al. 2011a; Bower et al. 2012; Clynes et al. 2013; Li et al. 2019). The loss of ATRX is variable depending on the cancer subtype although the majority of ALT positive sarcomas are null for ATRX expression (64%, table 1.2). In contrast, PanNETs, have an equal proportion of ATRX and DAXX mutations which appear to be mutually exclusive in the majority of cases (Kim et al. 2017; Singhi et al. 2017).

In vitro, the loss of ATRX in normal cells does not affect the phenotype as seen in hybrid cultures formed of normal and ALT positive cells which undergo telomere erosion as opposed

to elongation thus suggesting a repression of ALT in normal cells (Perrem et al. 1999). Similarly, when telomerase positive cells are co-cultured with ALT positive cells, the ALT phenotype disappears indicating that telomerase represses the ALT mechanism (Perrem et al. 1999). In addition, loss of ATRX alone in telomerase positive cells does not affect the telomere length, further suggesting maintenance of telomere length through telomerase. However, the appearance of APBs and C-circles can be observed thus suggesting these are required early in the process of ALT upregulation and initiation but other underlying mechanisms are required for long-term survival with ALT (Napier et al. 2015; Brosnan-Cashman et al. 2018). In contrast, lack of ATRX appears to facilitate escape from crisis by activation of the ALT mechanism in fibroblast cell lines therefore providing an interesting target to gain a better understanding in the mechanism of ALT initiation and survival (Napier et al. 2015).

Therefore, knocking out ATRX whilst inducing a telomere-driven crisis, by abrogation of TP53 and Rb, in primary human fibroblast will provide an interesting model for understanding the transition from normal to malignant cell growth, and potentially generate clones that upregulate ALT for survival to understand the mechanism of initiation.

3.3 Aims

The overall hypothesis for this chapter is as follows: the loss of ATRX combined with telomere stress is sufficient to induce replicative immortality by upregulating and maintaining the ALT mechanism in primary human fibroblasts. Throughout this chapter, primary human HCA2 fibroblasts were used for all experiments taken from a male donor.

The aims of this chapter are to:

- Establish if the loss of ATRX combined with telomere stress results in replicative immortality in primary fibroblasts in which TP53 function has been abrogated
- Assess if the loss alone of ATRX is sufficient to initiate a telomere-driven crisis in WT cells, with functional TP53, and subsequently initiating the ALT mechanism
- Analyse escaping clones, through C-circle and TRAP assays, and confirm the TMM used for replicative immortality
- Compare telomere dynamics between escaping clones and controls to reveal any characteristics related to ALT
- Compare fusion profiles between escapees and control clones and relate to the STELA profiles to establish any differences in the frequency and spectrum of fusion events
- Evaluate the involvement of other mechanisms, including the cGAS-STING pathway, in the upregulation and maintenance of ALT

3.4 Results

3.4.1 Analysis of HCA2 WT and HCA2$^{HPV\ E6E7}$ parental cell lines

The role of ATRX in suppressing the ability of primary human fibroblast to escape a telomere-driven crisis was examined using HCA2 primary human foreskin fibroblasts. These cells do not express telomerase and are subject to telomeric end-replication losses during each cell cycle and therefore, exhibit a finite telomere-dependent replicative lifespan (Wyllie et al. 2000). WT cells were taken into culture at PD 27 (approximately half their replicative lifespan) and were expected to divide until short telomeres triggered a TP53/Rb cell cycle arrest which in turn, elicits replicative senescence after approximately 30 PD (60 PD overall) (figure 3.1 A: green curve. NB t0 = 27 PD). Senescent cells were defined as large, flat, mono-nucleated cells that showed no increase in cell number for a period of two weeks.

Alongside these cells, a proportion of WT cells were transfected with HPV E6E7 viral oncoproteins by Mrs Julia Grimstead, at PD 30.5, to abrogate TP53 and Rb function to enable cells to bypass the senescent point and induce a telomere-driven crisis. This enabled cells to undergo an additional 33 PDs compared to the WT cells (90 PD overall) at which point, cells entered a telomere-driven crisis as defined by a halt in expansion resulting from an equal cell death to cell growth rate (figure 3.1 A: blue curve. NB t0 = 30.5 PD). Crisis cells were defined as uneven, large, multi-nucleated cells surrounded by cell debris and precipitate.

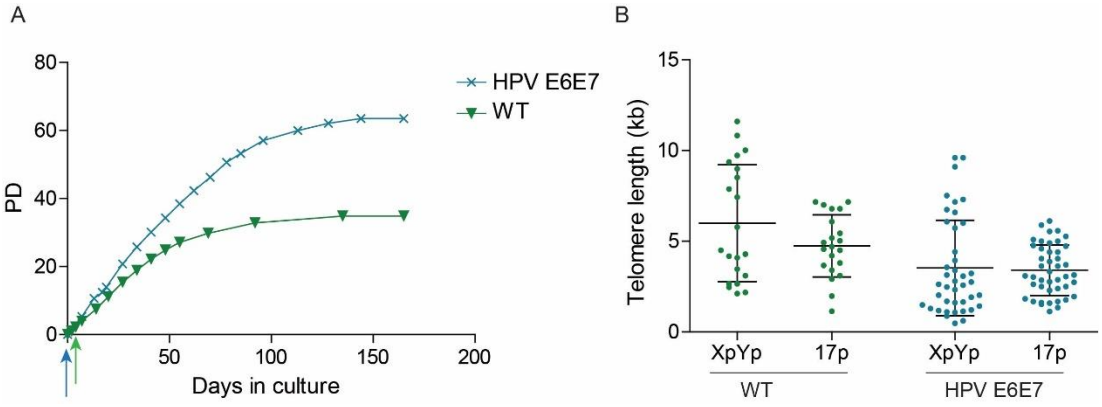


Figure 3.1: HCA2 cells have a limited replicative lifespan.
A) Growth curve, generated by Mrs Julia Grimstead, of the HCA2 WT parental cell line in green and the HCA2$^{HPV\ E6E7}$ parental cell line in blue. The green and blue arrows represent the WT (PD 3, 30 overall) and HPV E6E7 (PD 0, 30 overall) time-points respectively, used for STELA. Population doubling (PD) is depicted on the Y-axis B) Scatter plot representing the telomere length distributions determined with STELA at the 17p and XpYp chromosome ends in the HCA2 WT (green) and HCA2$^{HPV\ E6E7}$ (blue) parental cell lines (at PD point indicated by arrows on (A); SD error bars).

Telomere length profiles were determined with STELA at the 17p and XpYp chromosome ends at PD 30 overall in both parental cell lines (PD 3 for WT cells and PD 0 for HPV E6E7 cells pointed by the green and blue arrows respectively figure 3.1 A) to assess telomere dynamics. Despite analysing samples at the same PD point overall, the mean telomere lengths varied greatly between the parental cell lines (6 kb and 3.5 kb for XpYp and 4.7 kb and 3.4 kb for 17p for the WT and HPV E6E7 respectively) and were statistically significantly different (P value = 0.0033 and 0.0013 for 17p and XpYp respectively, following Mann-Whitney test: table 3.1). The telomeres in the HPV E6E7 expressing cells were shorter overall with the shortest telomeres reaching 600 bps as opposed to 2.1 kb in WT cells at the XpYp chromosome end, presumably due to the abrogation of TP53 and Rb enabling telomeres to shorten beyond the length at which senescence would normally be triggered. Despite observing different means between the two cell lines, the telomere distributions were similar with the presence of bimodal distributions consistent with two alleles at XpYp and one modal population at the 17p chromosome end consistent with a potential overlap of two similar allelic telomere length distributions (figure 3.1 B). In addition, standard deviations (SD) were comparable between the two cell lines and there was no statistical difference resulting from the Fligner test for homogeneity of variances when comparing chromosome ends between cell lines (P value = 0.418 and 0.103 for 17p and XpYp respectively: table 3.1).

	17p	**XpYp**
Mann-Whitney U test	0.0033	0.0013
Fligner test for homogeneity of variance	0.418	0.103

Table 3.1: Statistical differences between the telomere lengths in the WT and HPV E6E7 parental cell lines at the 17p and XpYp chromosome ends.
Table showing the p-values resulting from Mann-Whitney tests and Fligner tests for homogeneity of variances when comparing the telomere lengths in the WT and HPV E6E7 parental cell lines at the 17p and XpYp chromosome ends**.** Statistical differences are highlighted in red ($P < 0.05$).

3.4.2 HCA2 WT and HPV E6E7 control clones

3.4.2.1 HCA2 WT control clones characteristics

As experimental controls, HCA2 WT cells were plated at PD 31 at limited dilutions and 5 single cell clones were picked after 10 PDs (mean range: 6.5-13.5 PDs). All clones stopped dividing and entered senescence after approximately 23 PDs (mean range: 16.32 – 30.51 PDs) (figure 3.2 A). Telomeres at the XpYp and 17p chromosome ends eroded as a function of cell division until the shortest telomeres triggered replicative senescence consistent with previous

findings , notably with the XpYp lower allele reaching 1.62 kb in clone 5 (figure 3.2 B) (Baird et al. 2003). The mean telomere lengths for all single cell clones were determined at the last available sampling point (mean of 23 PDs following isolation of single cell clones) and were consistently shorter at the 17p chromosome end (mean of 3.02 kb) than the originating parental culture (4.74 kb) at PD 0, consistent with telomere erosion prior to the onset of senescence (figure 3.2 C; NB: Parental cell line sampling point at t=0 on graph A). In addition, a statistical difference was observed upon comparison of combined telomere lengths of control clones with the parental telomere lengths (P value <0.0001 following Mann-Whitney test). Despite the presence of long telomeres at the XpYp chromosome end, shorter telomere lengths (mean of 5.80 kb) compared to the parental (6.00 kb) were recorded in clones 1, 3, 4 and 5, although not statistically different (p value = 0.11 following Mann-Whitney test), consistent once more with telomere erosion and the triggering of replicative senescence.

Single molecule telomere fusion analysis of the unique XpYp and 17p telomeres, together with a family of telomeres with homology to the 21q chromosome end was undertaken to determine the base line frequency events in the HCA2 WT control clones (Letsolo et al. 2010). Consistent with previous data in young and senescent fibroblast populations, no telomere fusion events were detected in these cells (figure 3.2 D) (Capper et al. 2007).

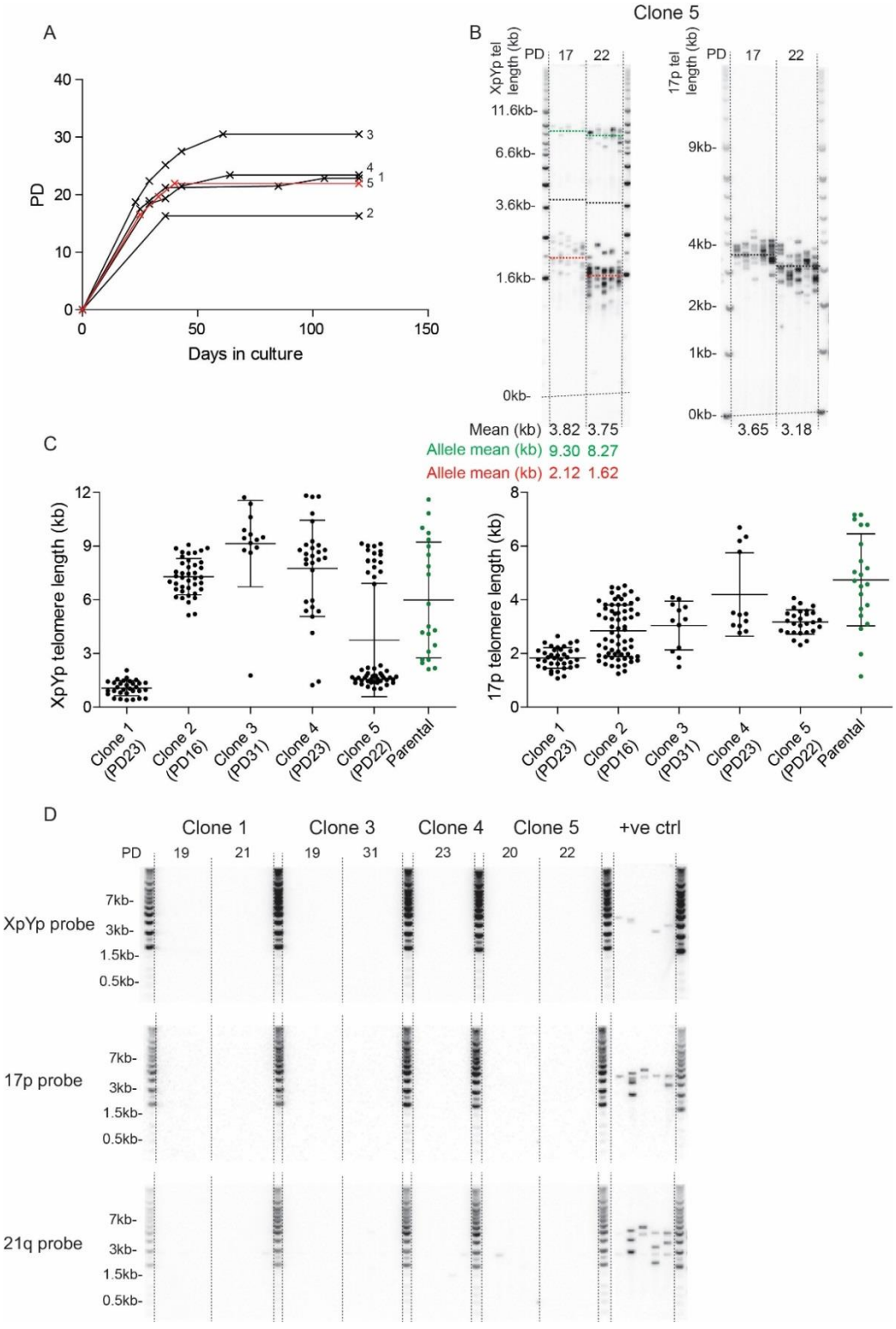


Figure 3.2: HCA2 WT control clones undergo replicative senescence.
A) Growth curves of the 5 single cell control clones with the PD depicted on the Y axis. In red, clone 5 for which the STELA profiles at the XpYp and 17p chromosome ends are shown in B) with the mean telomere lengths across the bottom also represented as a dotted line on the blot. In green the longer mean allelic telomere length and in red the shorter mean allelic telomere length. PDs stated across the top of the blot. C) Scatter plots representing the telomeres measured using STELA at the XpYp and 17p chromosome ends of the 5 single cell clones at the last sampling point and the WT parental cell line at PD 30 overall in green. D) XpYpM:17p6:21q1 fusion profiles for clones 1, 3, 4 and 5. Hybridisation probes are detailed on the left.

3.4.2.2 HCA2$^{HPV\ E6E7}$ control clones characteristics

As experimental controls, HCA2$^{HPV\ E6E7}$ cells were plated at PD 43 at limited dilutions and 6 single cell clones were isolated after 10 PDs (mean range: 8-12.5 PDs). Cells continued to divide for approximately 30 PDs (mean range: 15.04-41.7 PDs) until the onset of crisis followed by extensive cell death (figure 3.3 A). These cultures were passaged until all cells died (up to 230 days). Consistent with the telomere length profiles observed in the parental HCA2$^{HPV\ E6E7}$ cells, the XpYp telomeres of the clonal populations were long in all clones (mean of 4.92 kb), whilst the 17p telomeres were short (mean of 2.04 kb); as observed in clone 4 (figure 3.3 B). The mean telomere lengths for all single cell clones were determined at the last available sampling point (mean of 27 PDs after isolation of single cell clones) and were consistently shorter at the 17p chromosome end than the originating parental culture (3.40 kb) consistent with telomere erosion (figure 3.3 C; NB: Parental cell line sampling point 13 PDs prior to t=0 on graph A). A statistical difference was observed upon comparison of combined telomere lengths of control clones with the parental telomere lengths (P value <0.0001 following Mann-Whitney test). The lower allele at the XpYp chromosome end is overall shorter than in the parental further showing telomere erosion, despite no statistical difference (p value = 0.084 following Mann-Whitney test). These data indicate that the isolation of single cell clones is skewed towards clones with longer telomeres potentially due to the clones with short telomeres entering crisis early; before the point at which single cell clones were picked.

Single-molecule telomere fusion analysis was undertaken at PD points both prior to, and during, crisis. Fusions involving the XpYp telomere were uncommon across all clones analysed, whilst 17p fusions were more common (19 events overall at the latest sampling point), which coincided with the short 17p telomeres observed on the STELA blots and was consistent with previous observations (Capper et al. 2007; Letsolo et al. 2010) (figure 3.3 C and D).

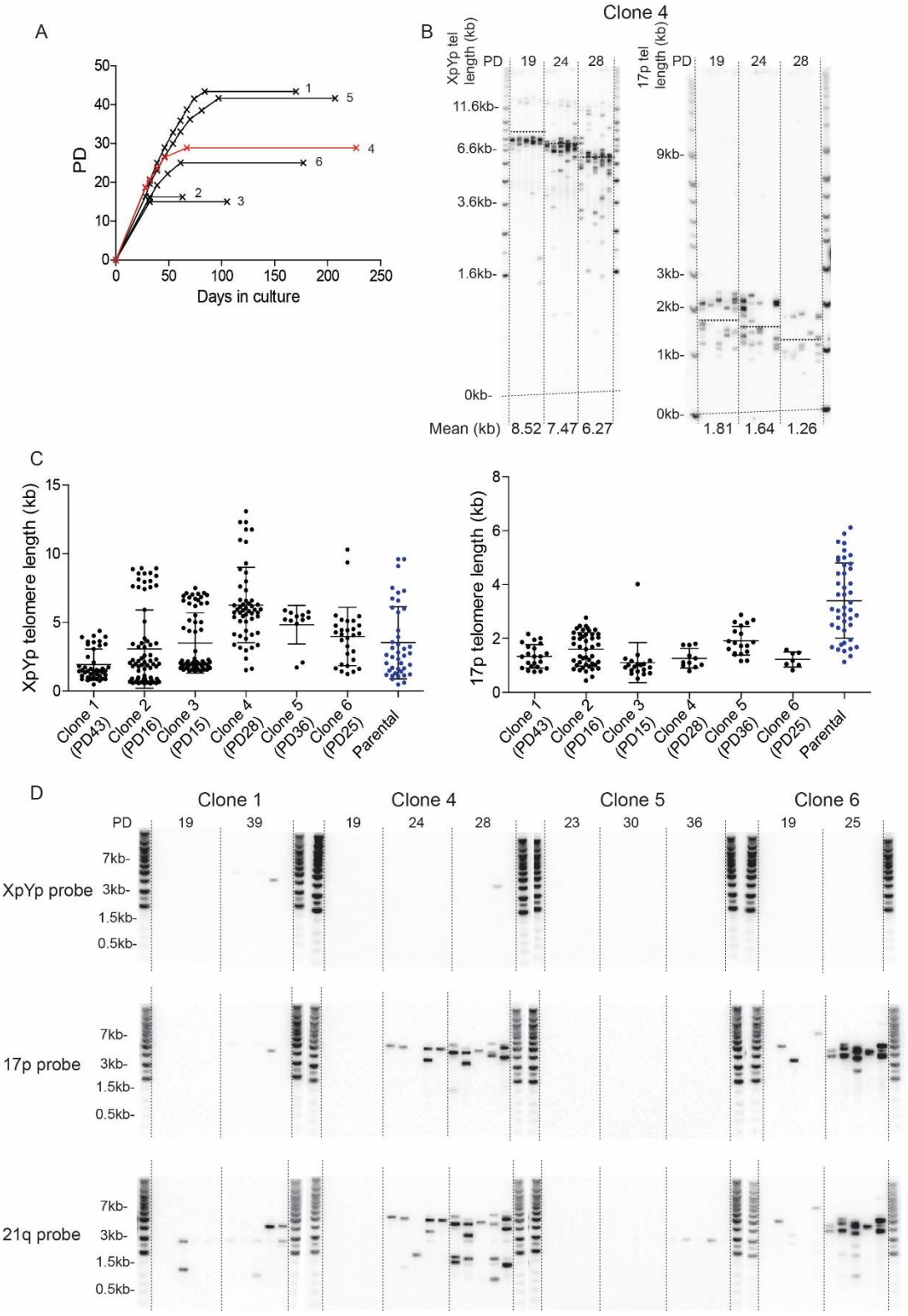


Figure 3.3: HCA2$^{HPV\ E6E7}$ control clones undergo crisis with short telomeres and fusions.
A) Growth curves of the 6 single cell control clones picked with the PD depicted on the Y-axis. In red clone 4 for which STELA profiles at the XpYp and 17p chromosome ends are shown in B) with the mean telomere lengths detailed across the bottom also represented as dotted lines on the blot. PDs stated across the top of the blot. C) Scatter plots representing the telomeres measured with STELA at the XpYp and 17p chromosome ends of the 6 single cells clones at the last sampling point during crisis with the HPV E6E7 parental telomere lengths at 13 PD prior to t0 in blue. D) XpYpM:17p6:21q1 fusion profiles of clones 1, 4, 5 and 6. Hybridisation probes are detailed on the left.

3.4.3 HCA2$^{HPV\ E6E7\ ATRX-/-}$ clones

3.4.3.1 Screening for mutant clones

In order to test whether primary fibroblasts could upregulate the ALT mechanism for survival in the absence of ATRX, HCA2$^{HPV\ E6E7}$ cells were nucleofected at PD 44.5 with an ATRX targeted CRISPR/Cas9 construct, provided by Prof. Eric Hendrickson (University of Minnesota, USA), containing a GFP tag for selection of transfected cells (figure 3.4 A) (Napier et al. 2015). These cells were then sorted by flow cytometry according to GFP intensity 24 hours after nucleofection (figure 3.4 B). To minimise background signal and reduce false positives, the selection gates were stringently applied to collect the strongest GFP intensity cells. This setup was consequently used in subsequent ATRX targeted knock out experiments using GFP to ensure consistency. A high percentage of cells were strongly positive for GFP and 118,000 cells were recovered from the cell sorting. Cells were then plated at limited dilutions (1:100, 1:250 and 1:500) in order to isolate single cell clones, consequently 52 clones were picked at a mean of 8.9 PDs after cell sorting (mean range: 6-12.5 PDs).

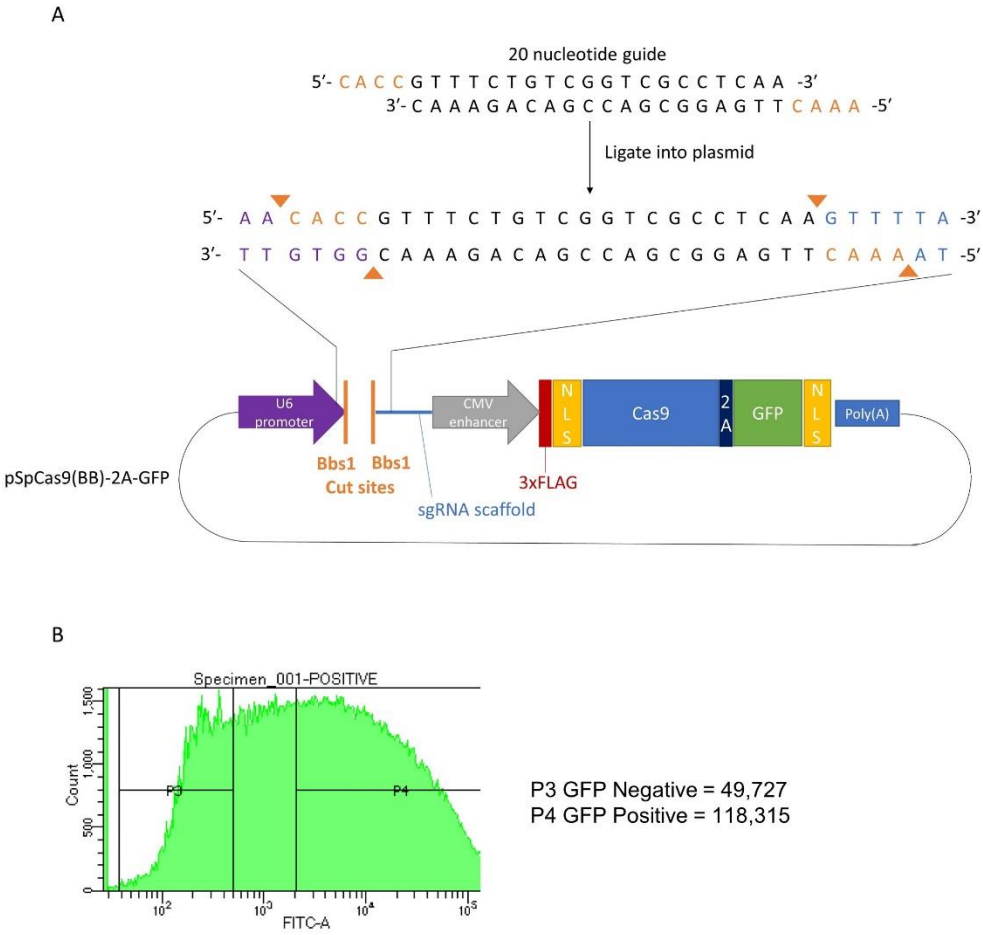


Figure 3.4: Schematic representation of the pSpCas9(BB)-2A-GFP vector containing the CRISPR targeting ATRX.
A) A 20 nucleotide guide sequence, designed to target exon 9 of the human *ATRX* gene, was ligated into the vector using Bbs1 restriction cut sites. B) Sorting of CRISPR targeted cells using GFP signal with P3 GFP negative cells (discarded) and P4 GFP positive cells (collected cells).

The *ATRX* gene is located on the X chromosome and due to HCA2 cells originating from a male donor, only one allele was required to be targeted and the CRISPR vector was designed to target exon 9 (figure 3.5 A and B). All clones were then screened to assess which had a successfully mutated *ATRX* gene. To do this, a 1 kb sequence from the *ATRX* gene, which included the CRISPR target site that contained a Sml1 restriction enzyme cut site, was amplified by PCR (figure 3.5 C and 3.6 A). 8 µl of the PCR product were digested with Sml1 restriction enzyme and the results were as follows: a single undigested 1 kb band signified a potential mutation at the cut site that removed the Sml1 restriction site; 2 bands measured at approximately 0.3 kb and 0.7 kb indicated an intact uncut *ATRX* gene (figure 3.6 A). Amplicons of putative mutant clones were verified by Sanger sequencing, using the AX9SeqF primer (Materials and Methods 2.4.1 Oligonucleotides and figure 3.5 C).

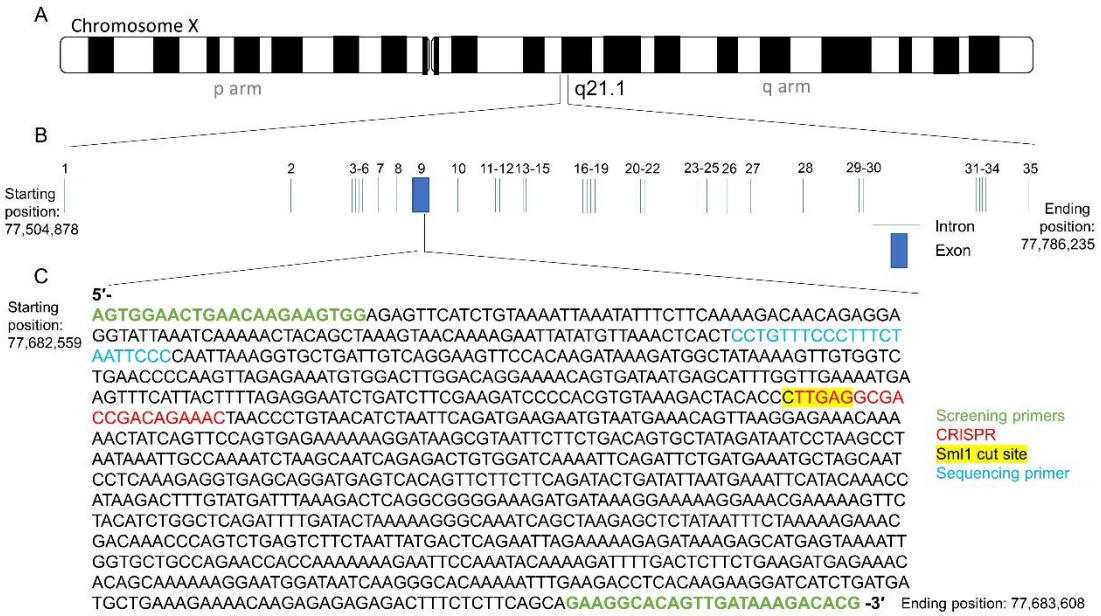
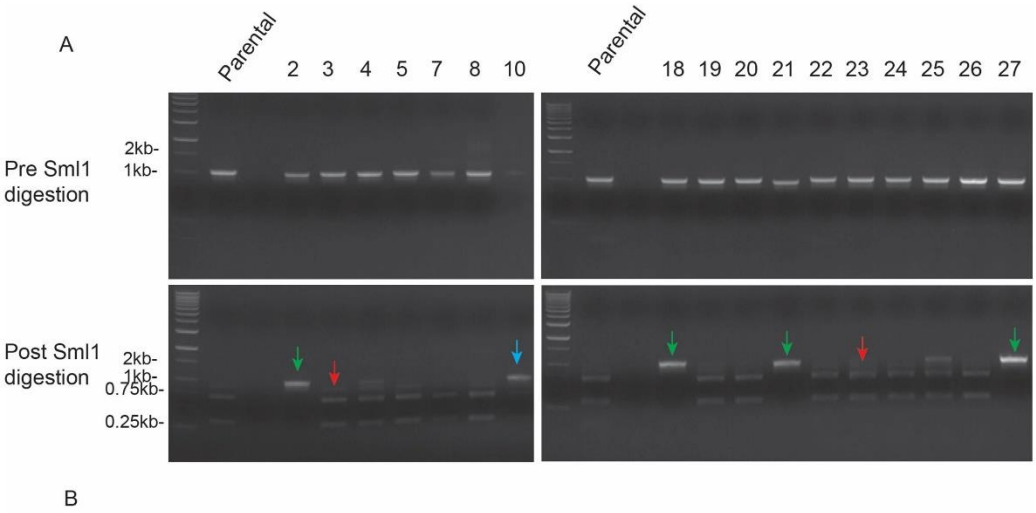


Figure 3.5: Schematic representation of the *ATRX* gene and CRISPR target site.
A) Representation of the X chromosome highlighting the position of the *ATRX* gene on the q arm (q21.1). B) Representation of the *ATRX* gene with the 35 exons represented as blue boxes with exon 9 being the largest containing the CRISPR target site. C) The 1 kb targeted sequence of the *ATRX* gene amplified by PCR with the PCR primers in green, the sequencing primer in blue, the CRISPR target site in red and the Sml1 cut site highlighted in yellow.

Of the 52 clones picked, 12 (1, 2, 10, 11, 18, 21, 27, 28, 37, 42, 44 and 49) showed a mutation at the Sml1 cut site and were sequence verified (figure 3.6). Following sequencing, 3 (10, 11 and 42) had an intact cut site and therefore an intact *ATRX* gene (for example, clone 10 figure 3.6 B), however, these were nonetheless cultured as controls to assess any off-target effects of CRISPR targeted technology on the ability of these cells to escape from crisis. The remaining 9 clones (1, 2, 18, 21, 27, 28, 37, 44 and 49) all showed mutations at the cut site indicating a high success rate of ATRX targeting (9/52 clones, 17%). This rate could be increased by applying more stringent GFP selection gates as the number of positive cells (118,000) was diluted in order to pick single cell clones and therefore many cells were dismissed during this process. The mutations detected ranged from a 2 base pair deletion (clone 18 figure 3.6 B) to a 16 base pair deletion including the entire cut site (clone 2 figure 3.6 B), as well as insertions and substitutions of single base pairs, always within the Sml1 cut site. These 12 clones were therefore continually passaged in culture to crisis and monitored for changes in cell morphology, telomere dynamics and fusion spectrum.



Clone 2

AGTTCCACAAGATAAAGATGGCTATAAAAGTTGTGGTCTGAACCCCAAGTTAGAGAAAT
GTGGACTTGGACAGGAAAACAGTGATAATGAGCATTTGGTTGAAAATGAAGTTTCATTA
CTTTTAGAGGAATCTGATCTTCGAAGATCCCCACGTGTAAAGAC - - - - - - - - - - - - - - -
CCGACAGAAACTAACCCTGTAA

Clone 10

AGTTCCACAAGATAAAGATGGCTATAAAAGTTGTGGTCTGAACCCCAAGTTAGAGAAAT
GTGGACTTGGACAGGAAAACAGTGATAATGAGCATTTGGTTGAAAATGAAGTTTCATTA
CTTTTAGAGGAATCTGATCTTCGAAGATCCCCACGTGTAAAGACTACACCCTTGAGGCG
ACCGACAGAAACTAACCCTGTAA

Clone 18

AGTTCCACAAGATAAAGATGGCTATAAAAGTTGTGGTCTGAACCCCAAGTTAGAGAAAT
GTGGACTTGGACAGGAAAACAGTGATAATGAGCATTTGGTTGAAAATGAAGTTTCATTA
CTTTTAGAGGAATCTGATCTTCGAAGATCCCCACGTGTAAAGACTACACCCTT- -GGCG
ACCGACAGAAACTAACCCTGTAA

Figure 3.6: CRISPR/Cas9 mediated mutation of ATRX in 17% of clones.
A) Gel electrophoresis of 1 kb ATRX amplicon including Sml1 CRISPR target site pre (top) and post (bottom) Sml1 digestion with green arrows showing mutants, blue arrow showing false positive mutants and red arrows showing examples of WT unmutated clones. B) Examples of Sanger sequencing data from sequence verified clones with the Sml1 cut site highlighted in yellow and hyphens indicating deletions within the sequence when compared to the WT ATRX sequence.

3.4.3.2 Clonal growth and ATRX protein expression

All clones were continuously cultured for a range of time (125 days to 250 days) to monitor changes in growth and to assess whether the loss of ATRX enabled these cells to escape crisis. Clones 10, 11 and 42, that displayed an intact *ATRX* gene, died after the onset of crisis. In addition, clones 1, 2 and 44 died following a prolonged crisis phase (figure 3.7 A). In contrast, clones 18, 21, 27, 28, 37 and 49 all transited through crisis and escaped. These all remained in crisis for varying amounts of time: a slow-down in growth was observed in clone 21 followed by a halt in cell division for 45 days as the cells were transiting through crisis (green

curve figure 3.7 B); whilst clone 37 remained in crisis for 100 days before resuming normal cellular growth (red curve figure 3.7 B).

Prior to crisis, clones that died had a similar growth rate as that observed in the HPV E6E7 control clones (mean of 0.41 and 0.46 PD/day respectively). In contrast, escaping clones underwent a more rapid growth rate prior to crisis (mean of 0.59 PD/day). Upon escape from crisis, these returned to a slower rate of cell growth (0.31 PD/day on average) and were continued in culture for up to 18 PDs post-crisis, until there were no crisis cells that were apparent, and cells were passaged every 7 days (figure 3.7 B).

Western blot analysis was undertaken to evaluate the status of the ATRX protein (figure 3.7 C). Clones 10, 11 and 42 (which had no mutation in the *ATRX* gene) had comparable or reduced levels of ATRX expression compared to the parental $HCA2^{HPV\ E6E7}$ cells (figure 3.7 D). In addition, clones 1, 2 and 44, which had a mutated *ATRX* gene but did not escape, all displayed residual ATRX expression (figure 3.7 C and D). In stark contrast, all clones that successfully escaped crisis had no detectable ATRX protein expression. This implies that low levels of residual ATRX activity is sufficient to prevent cells from escaping crisis and achieving replicative immortality through the upregulation of ALT in these fibroblasts and, suggests that the initiation of ALT requires the complete loss of ATRX early in the process of initiation. From this point, mention of ATRX will be referring to the protein in this chapter unless stated otherwise.

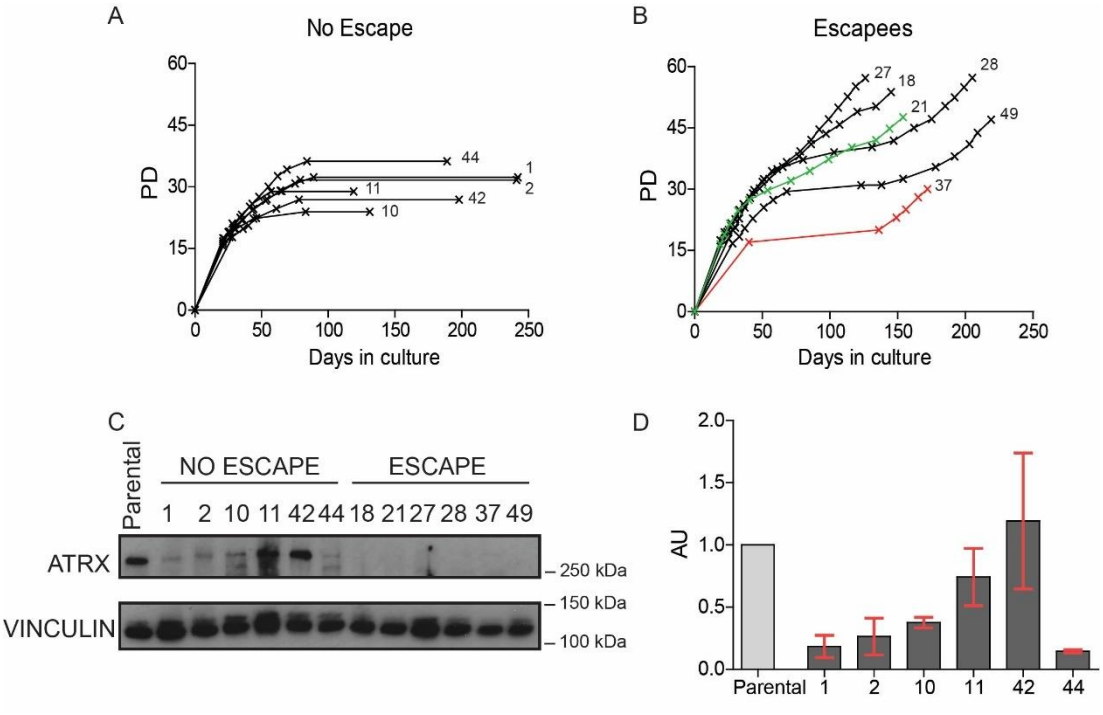


Figure 3.7: Complete ablation of ATRX results in cells escaping crisis.
Growth curves of the clones that failed to escape crisis (A) and successfully escaped crisis (B). PDs depicted on the Y-axis. C) Western blot showing levels of ATRX expression in the parental and 12 clones, vinculin expression was used as loading control. D) Western blot quantification of ATRX expression (ATRX : Vinculin ratio) relative to the parental, expressed in arbitrary unit (AU). Standard deviation used as error bars.

3.4.3.3 Confirming ALT positivity

The C-circle assay was developed to detect partially single stranded C-rich circular DNA present in the cytoplasm of cells called C-circles which are known as one of the quantifiable hallmarks of ALT (Henson et al. 2009). To establish the TMM used to escape crisis and achieve replicative immortality in these cells, all clones were subjected to the C-circle assay (figure 3.8 A). The HPV E6E7 control clones showed a comparable intensity to the HCA2$^{HPV\ E6E7}$ parental (figure 3.8 A and B) and were therefore considered negative for C-circles; the signal recorded was subsequently used as a baseline for the background signal. On this basis, the clones that did not escape crisis (1, 2, 10, 11, 42 and 44) were negative for C-circles (figure 3.8 A and B). In contrast, the escaping clones 18, 21, 27, 28 and 49 displayed no signal prior to crisis but this was followed by a gradual increase in C-circles as cells transited through crisis. Due to limited DNA available for clone 37, only a post-crisis sample was tested showing comparable intensity to other escaping clones (figure 3.8 A and B). This data confirms C-circle activity in 100% of clones that escaped crisis, consistent with the activation of the ALT mechanism in the absence of ATRX during a telomere-driven crisis.

Telomerase activity was assessed using the Telomerase Repeated Amplification Protocol (TRAP) assay to further establish whether escaping clones were capable of maintaining telomere length using the ALT mechanism, or whether a switch to telomerase and/or a co-existence of both mechanisms occurred. Pre- and post- crisis samples from clones that escaped crisis were tested for telomerase activity. Consistent with the parental HCA2$^{HPV\ E6E7}$ cells, all clones were negative for telomerase activity (figure 3.8 C). The absence of telomerase together with the presence of C-circles indicated that these clones extended their replicative capacity by utilising the ALT mechanism for telomere length maintenance.

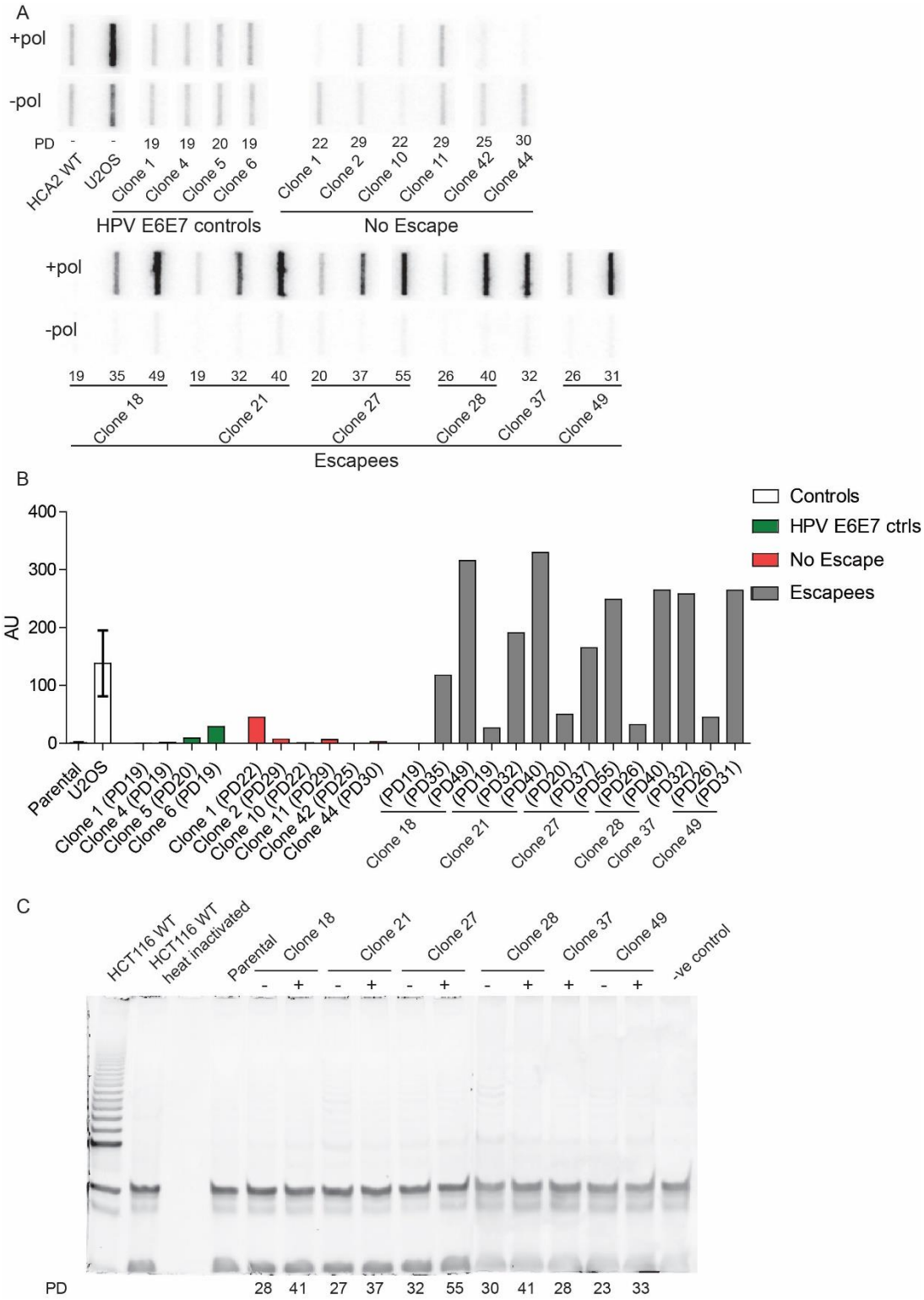


Figure 3.8: C-circles and absence of telomerase activity in clones that escape a telomere-driven crisis in the absence of ATRX is consistent with ALT.

A) '+' and '-' polymerase slot blot results of C-circle assay for HPV E6E7 control clones (1, 4, 5 and 6); clones that failed to escape crisis (1, 2, 10, 11, 42 and 44); and successfully escaped crisis (18, 21, 27, 28, 37 and 49). B) C-circle quantification with intensity of bands expressed in arbitrary unit (AU); HPV E6E7 controls in green, no escape clones in red, escape clones in grey and experimental controls in white ($HCA2^{HPV\ E6E7}$ parental cell line and U2OS ALT-positive control). SD used as error bars where possible. C) TRAP assay result before and after crisis (where possible) for clones that successfully immortalised with '-' representing before crisis and '+' representing after crisis. PDs across the bottom.

3.4.3.4 Telomere dynamics are affected by the loss of ATRX

To assess any changes and differences in telomere dynamics between escaping clones and clones that died, telomere length profiles were generated using STELA at the 17p and XpYp chromosome ends across multiple PD points. Clone 1, used as an example of a clone that did not escape crisis, showed a steady period of growth, 32 PDs, until the initiation of crisis (figure 3.9 A). Telomeres eroded as a function of cell division, as expected, and homogeneous distributions were observed at the chromosome ends studied consistent with the homogenous telomere length distributions observed in HPV E6E7 control clones (figure 3.9 B and figure 3.3 B and C). In contrast, clones that escaped crisis, for example clones 21 and 37 (figure 3.9 C), showed 2 different growth profiles with clone 21 entering crisis at approximately 27 PDs and transiting rapidly through crisis (approximately 45 days), whilst clone 37 entered crisis early (PD 17) and remained so for 100 days (figure 3.9 C). Nonetheless, telomere dynamics of these clones were similar, with both clones displaying heterogeneous telomere length distributions prior to the onset of crisis, in complete contrast to that observed in clones that failed to escape crisis, for example clone 1 (figure 3.9 B and D). Increased telomere length heterogeneity was also apparent post-crisis (figure 3.9 D). These heterogeneous telomere length profiles are consistent with that observed in cells that have activated the ALT mechanism. These data indicate that the loss of ATRX affects telomere length distribution before the full onset of crisis.

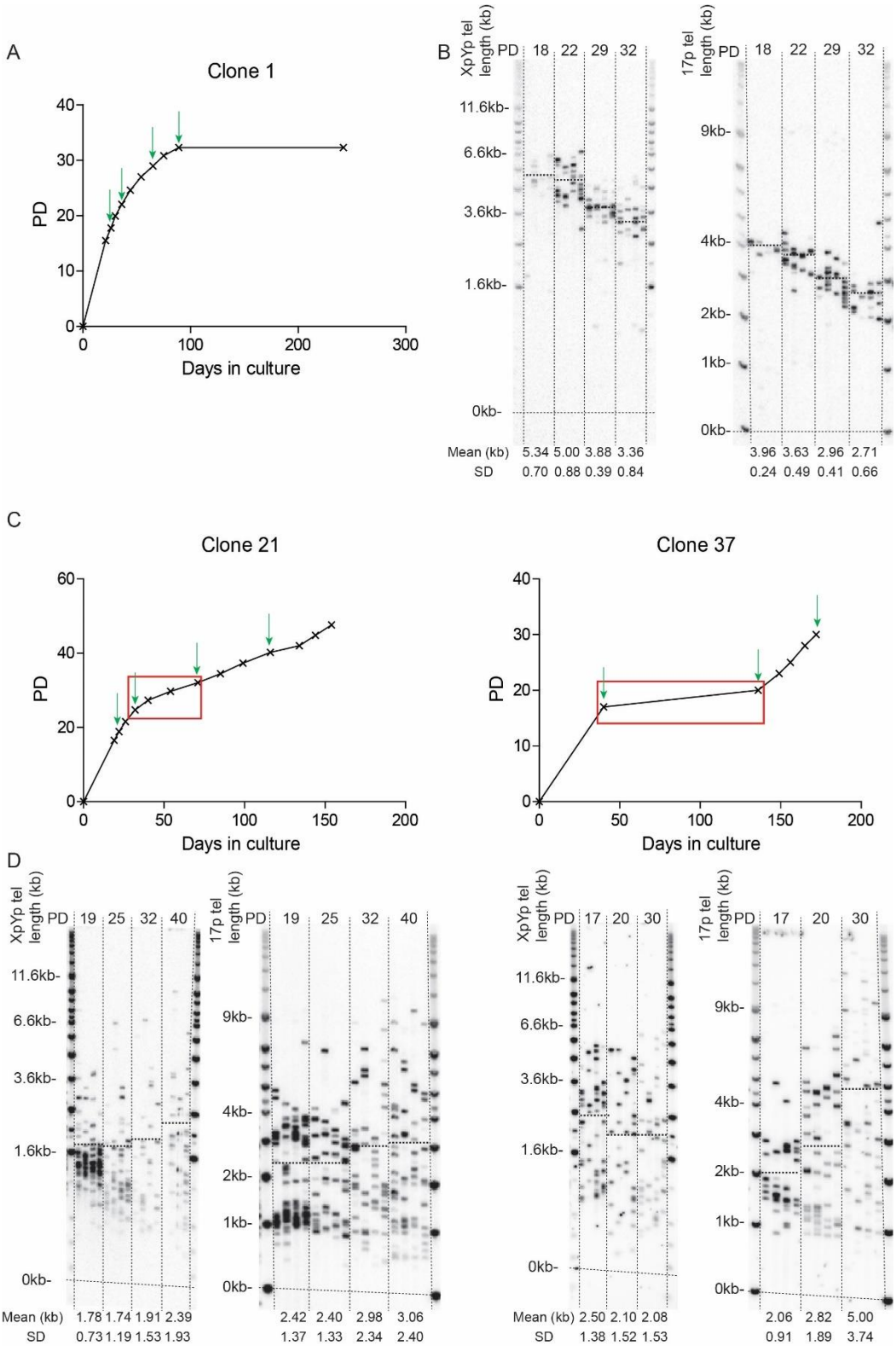


Figure 3.9: Loss of ATRX affects telomere length distributions before and after crisis.
A) Growth curve of clone 1 that did not escape crisis with PDs depicted on Y-axis. The green arrows represent time-points sampled for STELA at the 17p and XpYp chromosome ends showed in (B) with the mean telomere lengths across the bottom and represented as dotted lines on the blot. PDs across the top. C) Growth curves of clones 21 and 37 that successfully escaped crisis with the red rectangle highlighting the period of crisis and the green arrows representing the time-points used in the STELA profile at the 17p and XpYp chromosome ends showed in (D) with the SD and the mean telomere lengths across the bottom also represented as dotted lines on the blot. PDs across the top.

3.4.3.4.1 Increased heterogeneity before crisis in the absence of ATRX

To further address differences in heterogeneity between the controls, the clones that did not escape and the escaping clones, the telomere length distributions obtained from the first available time-point for all clones (18 PDs on average) were displayed as scatter plots with the standard deviation for both the 17p and XpYp chromosome ends. A single distribution was observed at the 17p chromosome end across all clones consistent with either the presence of a single 17p allele, or an overlap of two alleles in HCA2 cells. The controls and no escape clones all showed homogeneous telomere length distributions at 17p (mean SD = 0.623), whilst the escaping clones displayed a greater telomere length heterogeneity, with clone 28 presenting the biggest deviation from the mean (mean SD = 1.48; figure 3.10 A-C left). The standard deviations were grouped and plotted (figure 3.10 D left) and subjected to an unpaired t-test which showed statistically significant differences when comparing the escapees to the controls and the no escape clones (p = 0.0019 and 0.021 respectively). Furthermore, there was no statistical difference between the controls and the clones that failed to escape crisis (p = 0.58; figure 3.10 D left). This therefore suggests that the loss of ATRX leads to an increased heterogeneity prior to crisis, but only in the cells that were subsequently able to escape crisis.

In contrast, at the XpYp chromosome end, bimodal telomere length distributions were observed, and were consistent with allelic telomere length differentials, as previously observed (Baird et al. 2003). These bimodal distributions were only observed in the controls and the clones that failed to escape. In contrast, bimodal telomere length distributions were not apparent in the HCA2$^{\text{HPV E6E7 ATRX-/-}}$ clones that were capable of escaping crisis (figure 3.10 A-C right). The standard deviations for each group were plotted (figure 3.10 D right) and unpaired t-tests were carried out to compare each group, however it was clear that the presence of allelic telomere length distributions increased the heterogeneity of the control clones such that at the XpYp telomere, the controls appeared to be more heterogeneous (statistically different: p value = 0.041) than the HCA2$^{\text{HPV E6E7 ATRX-/-}}$ clones.

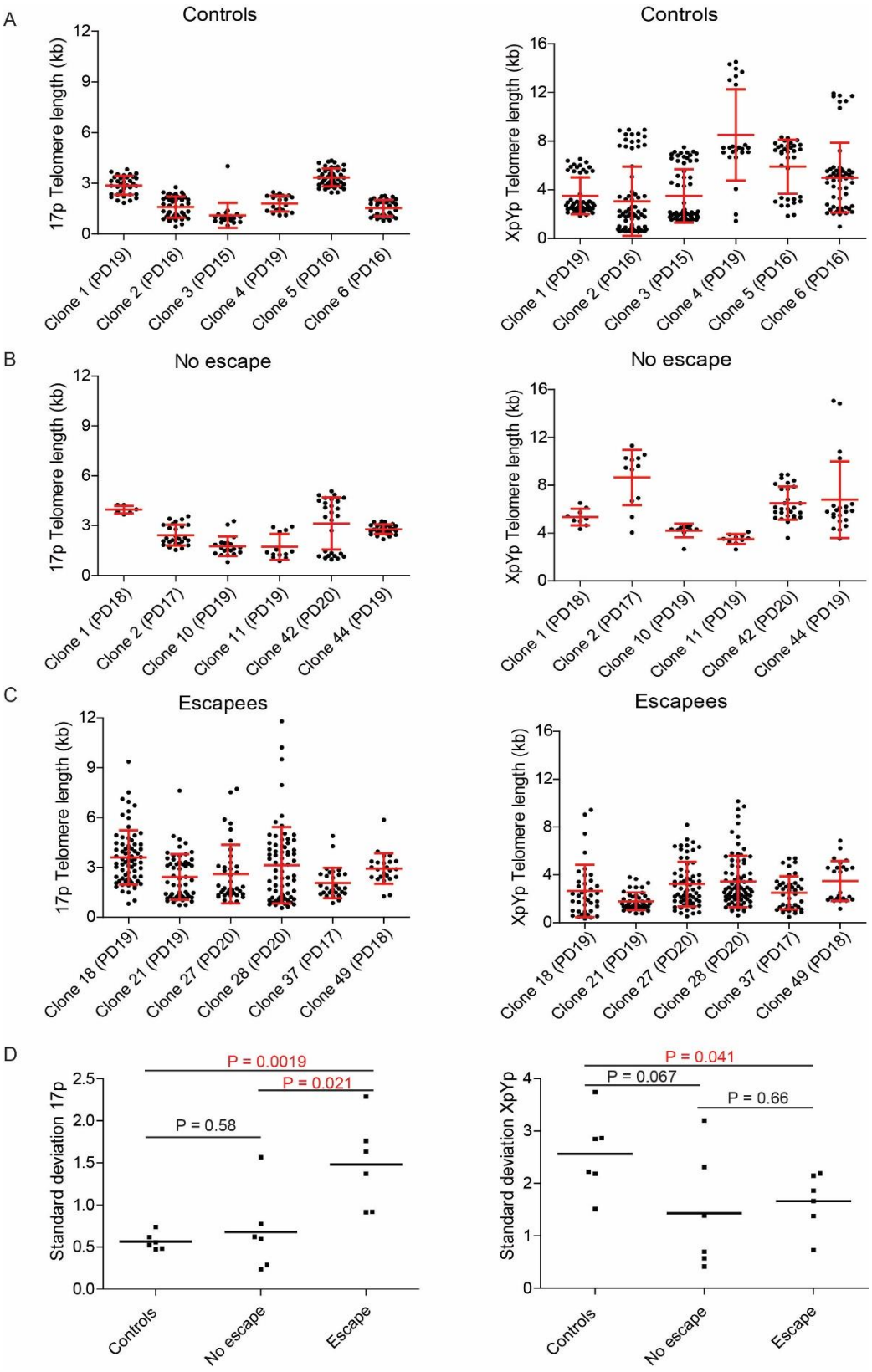


Figure 3.10: Heterogeneity before crisis in clones lacking functional ATRX.
Scatter plots of the telomere lengths determined using STELA at the 17p (left panel) and XpYp (right panel) telomeres in the HPV E6E7 controls (A), no escape clones (B) and escaping clones (C), with the SD as error bars. D) Scatter plot of the SDs for the controls, no escape clones and escaping clones for the 17p (left) and XpYp (right) chromosome ends with the P-value stated following an unpaired t-test; statistical differences are highlighted in red ($P < 0.05$).

3.4.3.4.2 Telomere length heterogeneity increases during the escape from crisis.

To further examine whether there was an increase in telomere length heterogeneity in the escaping clones, the 17p and XpYp telomere length profiles were displayed as scatter plots for each clone pre- and post- crisis (figure 3.11 A and B). At the 17p chromosome end, the telomere length heterogeneity consistently increased after crisis, whilst at the XpYp chromosome end, in most cases, this was maintained. Pre- and post-crisis samples for each clone were subjected to the Fligner test for homogeneity of variances and the p-values are stated above each clone in figure 3.11. At the 17p chromosome end, a statistically significant difference was observed for clones 21, 37 and 49. Meanwhile, clones 18, 27 and 28 showed no statistically significant differences in heterogeneity however, telomere lengths in these clones were already heterogeneous before crisis in comparison to clones 37 and 49 (figure 3.11 A). Furthermore, a statistically significant difference was observed upon comparing combined pre- and post- crisis standard deviations using an unpaired t-test, thereby further underlining more heterogeneous 17p telomere lengths after crisis ($p = 0.017$; figure 3.11 C).

In contrast, at the XpYp chromosome end, all clones but 21 and 28 showed no statistically significant differences in telomere length heterogeneity pre- and post-crisis. Clone 28 showed a decrease in heterogeneity (figure 3.11 B). Furthermore, no statistical difference was noted when comparing combined pre- and post- standard deviations (p value = 0.76, following unpaired t-test) (figure 3.11 C). In addition, there was a statistical difference upon comparison of 17p and XpYp SDs post-crisis (p value = 0.048, following unpaired t-test). These data imply that the 17p telomere may be subjected to larger elongation events compared to XpYp telomeres and could explain the slower increase in heterogeneity at the XpYp chromosome end.

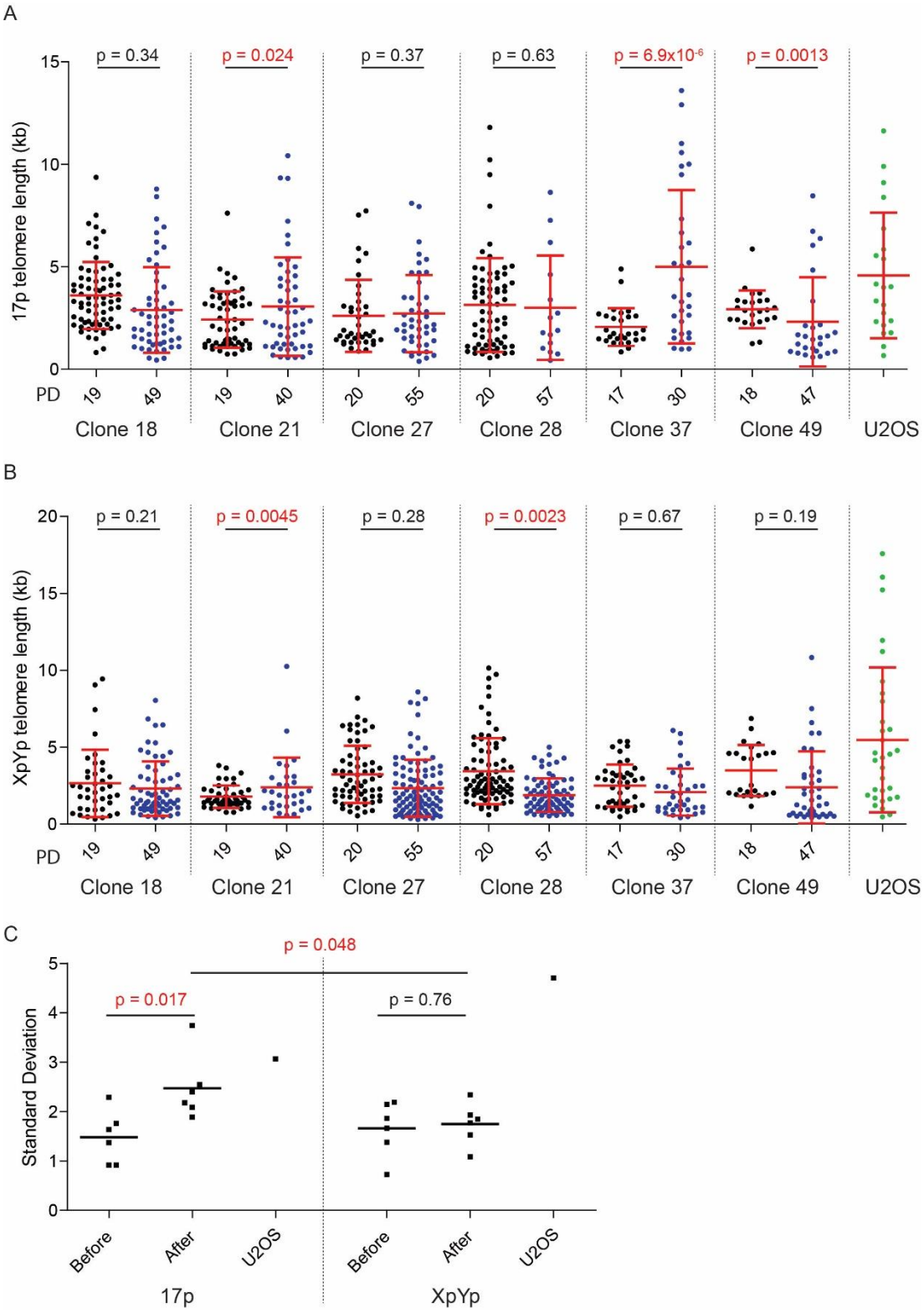


Figure 3.11: An increase in 17p telomere length heterogeneity after crisis in escaping clones. Scatter plots of the telomere lengths determined using STELA at the 17p (A) and XpYp (B) chromosome ends for the escaping clones pre- (black) and post- crisis (blue) and the U2OS ALT positive cancer cell line (green) with the SD as error bars. The p-values stated above samples resulting from Fligner's test for homogeneity of variances; statistical differences are highlighted in red ($P < 0.05$). C) Scatter plot of the standard deviations before and after crisis at the 17p (left) and XpYp (right) chromosome ends with the U2OS SD as an indication. P-values stated resulting from unpaired t-test; statistical differences are highlighted in red ($P < 0.05$).

The telomere length heterogeneity at both chromosome ends was compared to the ALT positive cancer cell line U2OS in all clones. Using the Fligner test for homogeneity of variances, overall, no difference was noted when combining all escaping clones against the U2OS at the 17p telomere, further highlighting the more heterogeneous telomere length profile at this chromosome end (P value = 0.097; table 3.2). Consistent with the previous observations, a strong statistical difference was observed at the XpYp chromosome end (P value = $1.1x10^{-9}$; table 3.2). All clones were then individually compared to the U2OS after crisis and once more, at the 17p chromosome end, half the clones showed no statistical differences, whilst the remaining three showed p-values greater than 0.01. In addition, all clones were statistically different to the U2OS at the XpYp chromosome end and in most cases the p-value was lower than 0.001.

	U2OS	
	17p	Xpyp
Clone 18	0.038	$4.4x10^{-5}$
Clone 21	0.13	0.00022
Clone 27	0.019	$1.7x10^{-6}$
Clone 28	0.24	$8.6x10^{-9}$
Clone 37	0.43	0.00019
Clone 49	0.030	0.0028
All	0.097	$1.1x10^{-9}$

Table 3.2: Heterogeneity at the 17p chromosome end consistent with U2OS ALT positive cell line. Table presenting the P-values following a Fligner test for homogeneity of variances between individual clones (left) and the U2OS (top) at the 17p and XpYp chromosome ends alongside the overall difference when combining all escaping clones post-crisis and comparing to the U2OS (termed all). Statistical differences are highlighted in red ($P < 0.05$).

3.4.3.5 Increased fusion events in ALT clones

Telomere fusion events are readily detected in fibroblast clones undergoing crisis (Capper et al. 2007; Letsolo et al. 2010). To examine whether the absence of ATRX impacted telomere fusions in $HCA2^{HPV\ E6E7}$ fibroblast cells, fusion profiles were generated for both the escapees and the clones that failed to escape. All clones were tested for XpYp:17p:21q fusions at multiple time-points in sequential order by the process of stripping and re-probing the same fusion blot (figure 3.12 A). Furthermore, fusions that involved inter-allelic events were counted once, therefore identical bands that appear with two probes were counted once and attributed to the first probe that detected these. A series of fusions that appeared at the same length across multiple reactions were discounted as only unique events were counted.

Fusions involving the XpYp telomere were sparse in both the controls and clones that failed to escape, presumably due to the initial XpYp telomere length being comparatively long (figures 3.3 B-D; 3.9 B; 3.10 A and B; and 3.12 A). In contrast, fusions involving the XpYp telomere appear at the first sampling point in escaping clones, presumably due to their heterogeneous telomere length distributions with the shortest telomeres reaching just 700 bps (figure 3.9 D, 3.10 C and 3.12). In addition, 17p fusion events were observed in all clones, consistent with short telomeres observed in the control and no escape clones as well as heterogeneous telomere lengths in the escapees (figure 3.3 B-D; figure 3.9 B and D; 3.10 A-C; and 3.12 A). The length at which telomeres begin to fuse at the 17p chromosome end, approximately 1 kb, is consistent across all clones therefore suggesting that the loss of ATRX does not alter the length at which telomeres begin to fuse.

Fusion events were monitored in all escapees across time in culture. Clone 21, used as an example of an escapee (figure 3.12 A), showed a gradual increase in telomere fusions at all telomeres studied as the clone transited through crisis. This was followed by a gradual decrease as cells immortalised and escaped through initiation of the ALT mechanism. This pattern was consistently observed across all escapees and was consistent with observations made in the U2OS cell line for which no fusions were detected at all chromosome ends analysed (figure 3.12 A and B). This suggests that the fusions observed in the escaping clones are associated with telomere stress, rather than ALT, and the progressive decrease in fusion events is consistent with ALT upregulation and the maintenance of telomere length for survival.

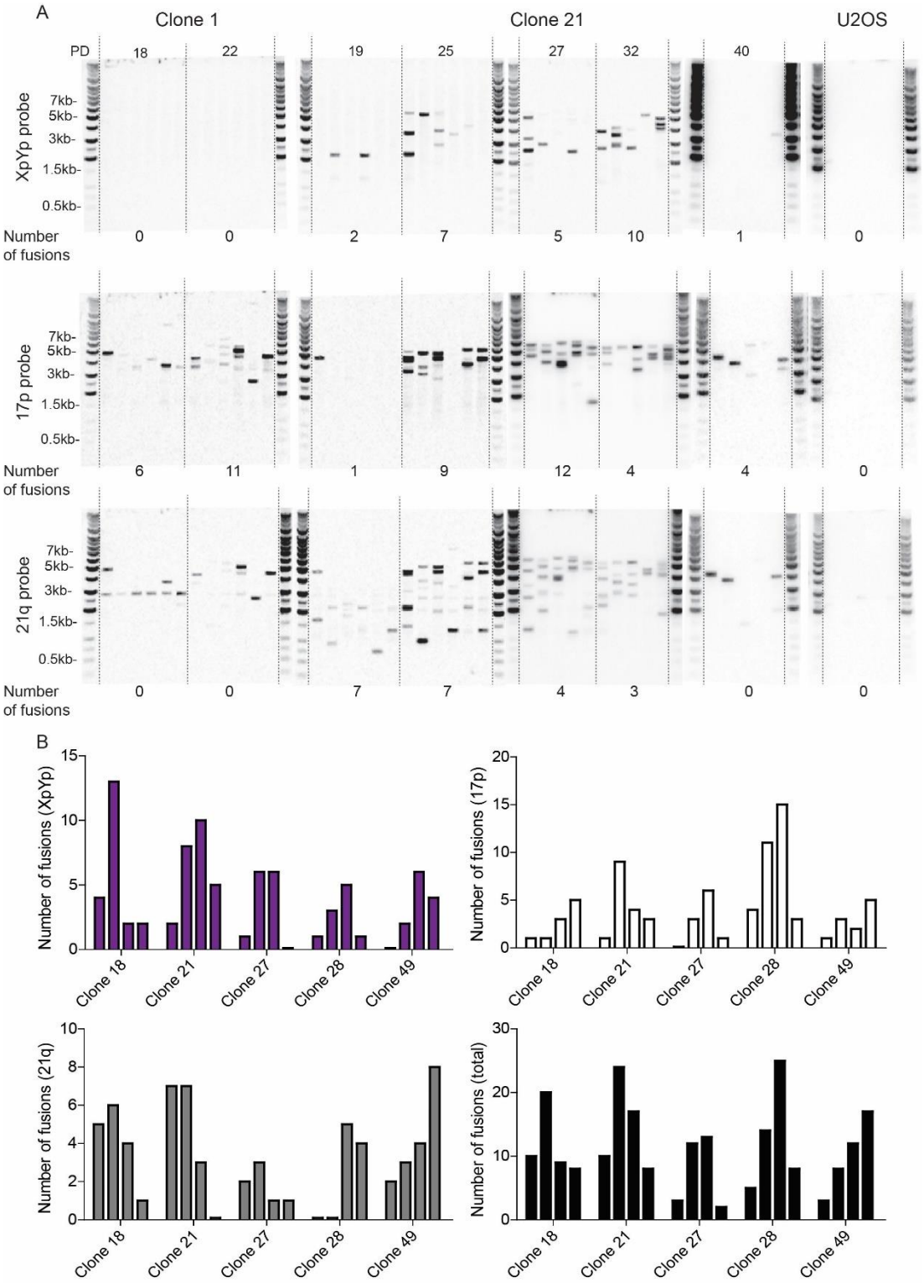


Figure 3.12: Fusion profiles reveal an increase of end-to-end fusions in escapees during crisis.
A) Example of fusion profiles for clone 1 (no escape) and clone 21 (escape) across multiple PD points (detailed across the top) and U2OS as an ALT positive control. Blots were serially probed and stripped with the number of fusions stated below the blots, only unique events were counted and events appearing with two probes were counted once. Hybridisation probes are detailed on the left. B) Graphs showing the number of XpYp (purple), 17p (white), 21q (grey) and total (black) fusion events as escapees are transiting through crisis and immortalising. (Number of diploid genome equivalents analysed = $2x10^4$).

It appeared that fusions occurred earlier and at higher frequency in the absence of ATRX. To compare the frequencies and fusion spectrums between all three groups, fusions were counted at the same point in the replicative lifespans across all clones (approximately 5 PDs prior to the slow-down in growth; figure 3.13: orange arrows representing the sample analysed for each clones). The proportion of fusions was measured by calculating the percentage of XpYp, 17p and 21q fusion events out of the total number of fusion events recorded for each group (controls, no escape and escape). Overall, the controls and the clones that failed to escape crisis showed similar fusion proportions (figure 3.14 A) and consistently showed no statistically significant differences when comparing the three chromosome ends studied upon submission to a Fisher's exact test (P values = 0.56, 1 and 1 for XpYp, 17p and 21q fusion proportions respectively). In contrast, escapees showed an even distribution of XpYp, 17p and 21q fusions (figure 3.14 A) and a statistically significant difference was observed in the proportion of 17p fusion events compared to the controls and no escape clones (p value = 0.011 and 0.0001 respectively following Fisher's exact test). Furthermore, there was a higher incidence of XpYp fusions in the escapees consistent with a more heterogeneous telomere length distribution and the presence of very short telomeres as opposed to the no escape clones (figure 3.9 B and D; 3.10 B and C; 3.12 B; and 3.14 A; table 3.3). In addition, there was an increase in fusions overall and in the mean frequency of fusions per escaping clones (10.8 versus 6.7 for the no escape and 3.3 for the controls; figure 3.14 B).

To further assess the nature (inter- or intra- chromosomal) of the fusion events observed, a PCR was repeated with single XpYp and 17p primers. No fusions were observed at the XpYp chromosome end, therefore, statistical analysis was carried out at the 17p chromosome end. The results showed an even proportion of intra-chromosomal fusions relative to the total number of fusions per group (figure 3.14 C; table 3.3) and the loss of ATRX therefore does not affect the rate of intra-chromosomal fusions but rather inter-chromosomal fusions as seen with the overall increase in fusion events in the escaping clones (figure 3.14 B).

		XpYp	17p	21q
All fusions	Controls vs Escape	0.16	0.011	0.32
	Control vs no escape	0.56	1	1
	No escape vs escape	0.0014	0.0001	0.10
Intra-chromosomal fusions	Controls vs Escape	-	0.65	-
	Control vs no escape	-	0.39	-
	No escape vs escape	-	0.78	-

Table 3.3: Higher rate of inter- but not intra-chromosomal fusions in escaping clones.
Table presenting the p-values resulting from Fisher's exact tests comparing the 3 groups against each other at each telomere end studied looking at all fusions and intra-chromosomal fusions only.

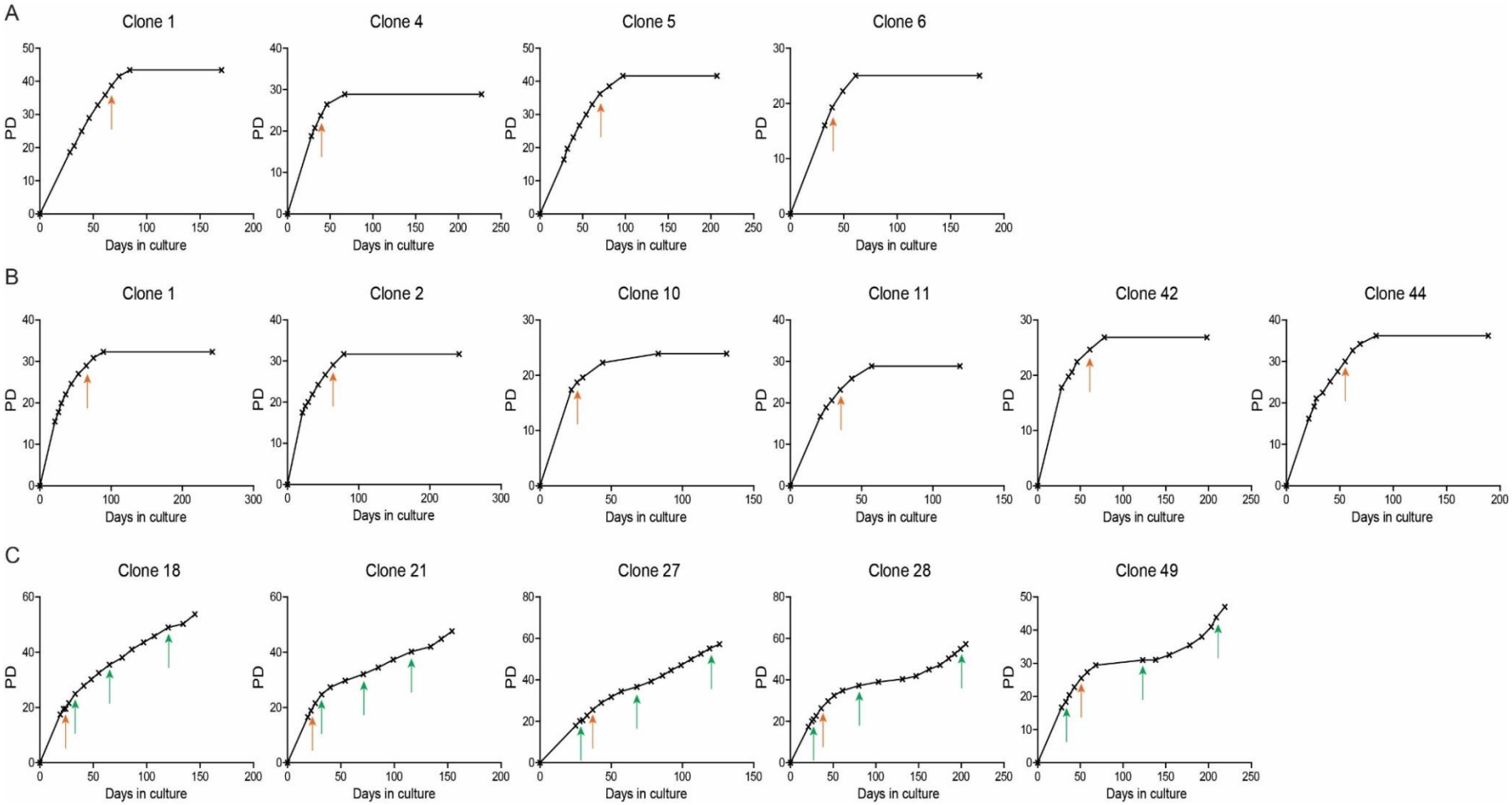


Figure 3.13: Growth curves of all clones analysed for fusion analysis.
A) HPV E6E7 control clones; B) no escape clones; C) and escapees with the orange arrows indicating the time-point used to compare fusion events across all clones and the green arrows the samples used for STELA analysis. PDs depicted on the Y-axis.

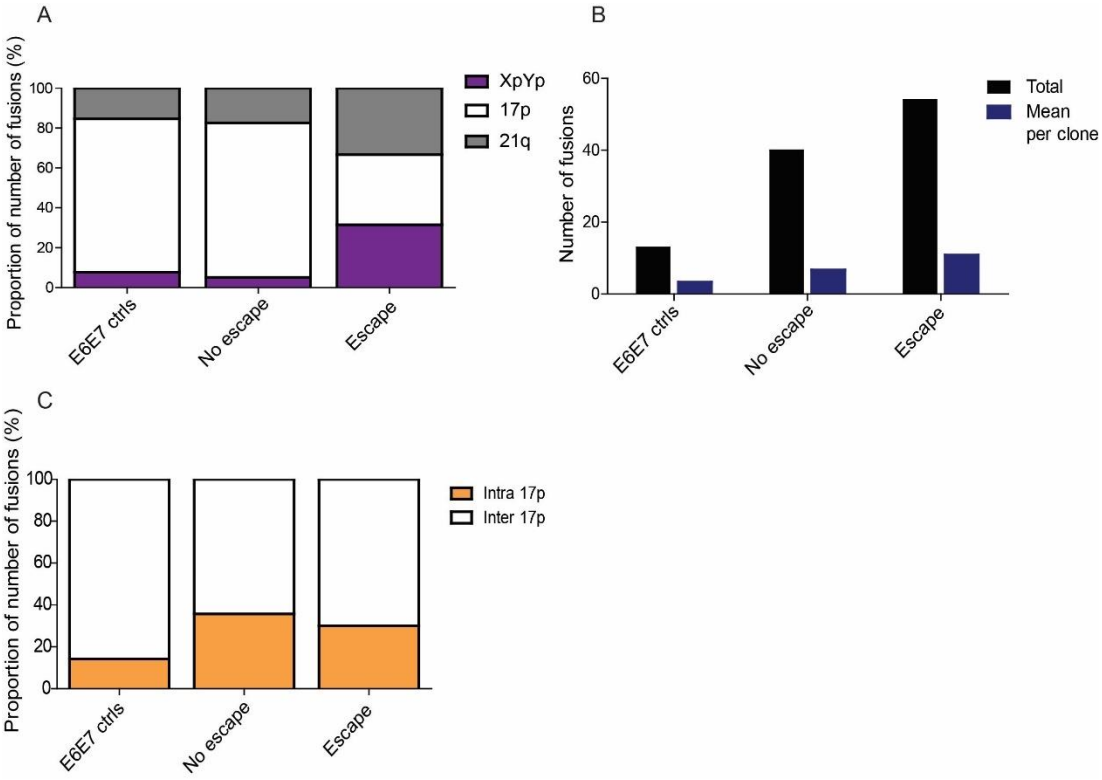


Figure 3.14: Altered proportion of fusion events in ALT clones.
A) Graph showing the proportion of specific telomere fusions for the controls, no escape and escaping clones (%). B) Graph showing the overall number of fusions per group (black) and the mean number of fusions per clone in each group (blue). C) Graph showing the proportion of 17p intra-chromosomal fusions (orange) and 17p inter-chromosomal fusions (white) for each group (%).

3.4.3.6 Impaired cGAS-STING pathway

The DNA sensing cGAS-STING pathway functions by recognising DNA in the cytoplasm of cells through cGAS binding which in turn releases the STING protein from the endoplasmic reticulum and promotes the phosphorylation of IRF3. IRF3 then migrates to the nucleus where it upregulates interferon-β expression which upregulates the innate immune response (Chen et al. 2016). Recently, the cGAS-STING pathway has been shown to be impaired in ALT positive cancer cell lines, such as U2OS, thus showing an impaired DNA sensing mechanism (Chen et al. 2017). This in turn enables the accumulation of C-circles in the cytoplasm of ALT-positive cells without inducing an immune response.

To assess whether the cGAS-STING pathway was impaired in escaping clones compared to clones that failed to immortalise, STING protein levels were determined by Western blot (figure 3.15 A). The clones that failed to escape crisis lacked C-circles and died, and thus STING levels were expected to be comparable to that observed in the parental cells. Expression levels varied between clones with a decrease in clones 1 and 2, an increase in clone 44 and

comparable expression in clone 11 to the parental. Furthermore, clones 10 and 42 showed no expression of the STING protein (figure 3.15 B). In comparison, the escaping clones 21, 28, 37 and 49 showed very little to no signal before and after crisis escape. Clone 18 appeared to have normal levels before crisis which disappear following escape suggesting an effect on STING activity in this clone. Clone 27 showed an increased protein expression before crisis which became comparable to the parental after crisis. The STING protein appeared to be consistently downregulated in the escaping clones (apart from clone 27) however, a similar observation was made in the clones that failed to escape (figure 3.15 B). The loss of the STING protein may be required for long-term survival with ALT to prevent the accumulation of C-circles from triggering the cGAS-STING pathway which in turn elicits a cell cycle arrest. Overall, there was no clear pattern of STING expression with respect to whether the clones escaped crisis or not, thus in this cell model, the loss of the STING protein is not a characteristic specific to the ALT mechanism in the absence of ATRX but may nonetheless facilitate cellular growth despite strong C-circle intensity.

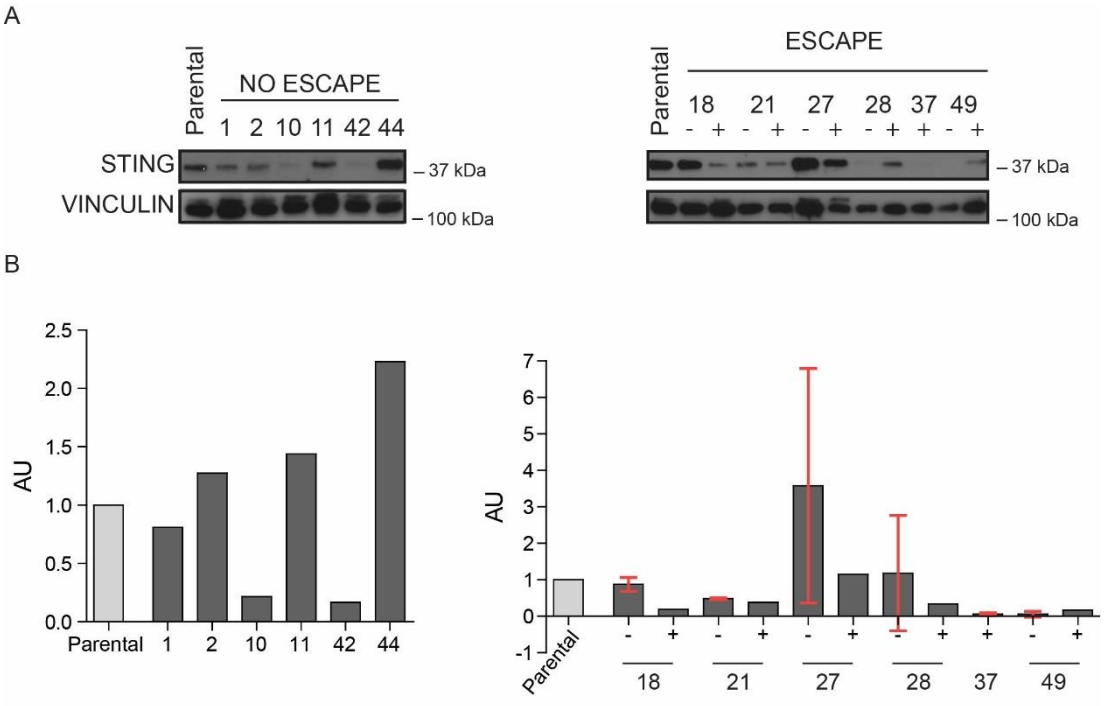


Figure 3.15: Decreased STING protein expression in ALT positive clones.
A) Western blot result showing levels of STING protein expression in the parental, clones that did not escape (1, 2, 10, 11, 42, 44) and clones that escaped (18,21, 27, 28, 37 and 49), vinculin expression was used as loading control. B) Western blot quantification of STING expression (STING : Vinculin ratio) relative to the parental expressed in arbitrary unit (AU). The '-' representing before crisis and '+' after crisis. Standard deviation used as error bars where possible.

3.4.4 ATRX knock out in WT cells

3.4.4.1 HCA2$^{ATRX-/-}$ WT single cell clones

It was clear that the complete absence of ATRX coupled with the onset of a telomere-driven crisis was sufficient to initiate the ALT process in cells that were abrogated for TP53 and Rb function. In contrast to the HCA2$^{HP\ VE6E7\ ATRX-/-}$ cells, WT cells have an intact TP53/Rb cell cycle checkpoint and thus the hypothesis that the onset of replicative senescence coupled with the loss of ATRX may lead these cells to gain replicative immortality by upregulating the ALT mechanism for survival was tested. HCA2 WT cells were nucleofected at PD 31 with the CRISPR vector in an attempt to knock out ATRX. Following nucleofection, an 80% survival rate of cells was observed which is an indicator of a low nucleofection efficiency. Cells were then sorted by flow cytometry according to GFP signal using the same settings as the previous experiment (figure 3.4 B) and the results are presented in figure 3.16 A. Low efficiency of the nucleofection (1.7%) was confirmed with most cells (98.3%) being GFP negative and only 2,000 cells being collected for single cell cloning. Nonetheless, 23 single cell clones were isolated approximately 8.8 PDs after cell sorting (mean range: 5-11 PDs) and 15 were screened for an ATRX mutation, whilst 8 senesced before a sample was taken. No clone exhibited a mutation in the *ATRX* gene and 5 clones were cultured to monitor any changes despite having a WT ATRX. The 5 clones (highlighted in red figure 3.16 B) were kept in culture for 120 days and a senescent phenotype was observed across all clones and no changes were visible until the cultures were intentionally terminated.

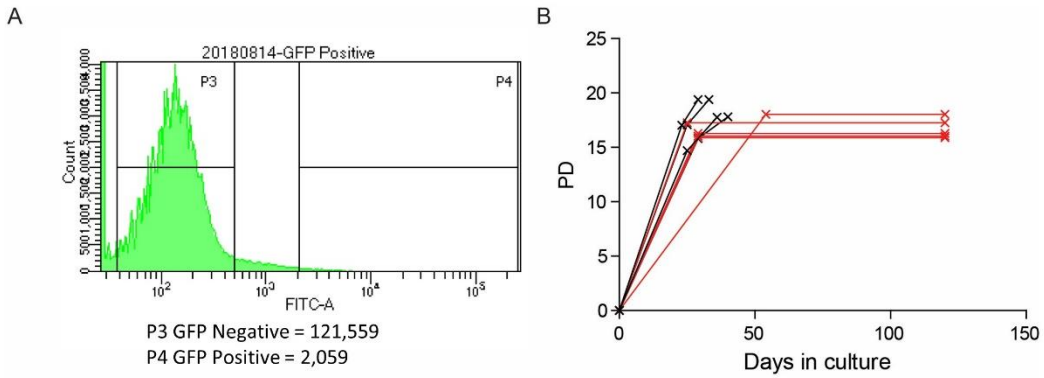


Figure 3.16: Low nucleofection efficiency and an absence of ATRX mutation in HCA2 WT cells resulted in all clones entering senescence.
A) Cell sorting result with P3 = GFP negative (discarded) and P4 = GFP positive (cells being collected). B) Growth curves of 9 clones that were targeted with the CRISPR and in red clones that were kept in culture long-term (n=5). PDs depicted on Y-axis.

3.4.4.2 HCA2$^{ATRX-/-}$ WT mixed population

It was considered that if the ATRX knockout could facilitate the induction of ALT and the escape from replicative senescence, then cells displaying an ATRX mutation may be selected for during the onset of senescence. Thus, following single cell cloning, the remaining cells that had been subjected to the ATRX CRISPR transfection and cell sorting, were pooled and transferred to a flask and kept in culture to investigate the behaviour of the polyclonal mixed population (MP) during senescence. In addition, a subset of cells taken from the MP were nucleofected at two later time-points, closer to senescence, to assess if a second ATRX-targeted CRISPR could yield HCA2 WT cells with an ATRX mutation. Firstly, the MP was nucleofected at PD 17 (MPN3, green arrow figure 3.17 A; growth curve 3.17 B) and secondly at PD 22 (MPN5, blue arrow figure 3.17 A; growth curve 3.17 B). Following nucleofection, cells were sorted once more by flow cytometry according to GFP signal and again, the nucleofection efficiency was low and very few cells were collected (figure 3.17 C and D). All three populations were kept in culture and screening for mutations revealed no evidence of an ATRX modification (data not shown). These cells were therefore expected to undergo approximately 30 PDs before senescing. These cultures all achieved a similar number of PDs before entering senescence as the WT parental control (figure 3.17 B). Similar telomere length distributions were observed throughout and this revealed an erosion of the telomeres with no evidence of the short telomere length distributions observed in crisis (figure 3.17 E and F). Moreover, the C-circle assay was negative for the three MP cultures confirming the lack of ALT initiation (figure 3.17 G).

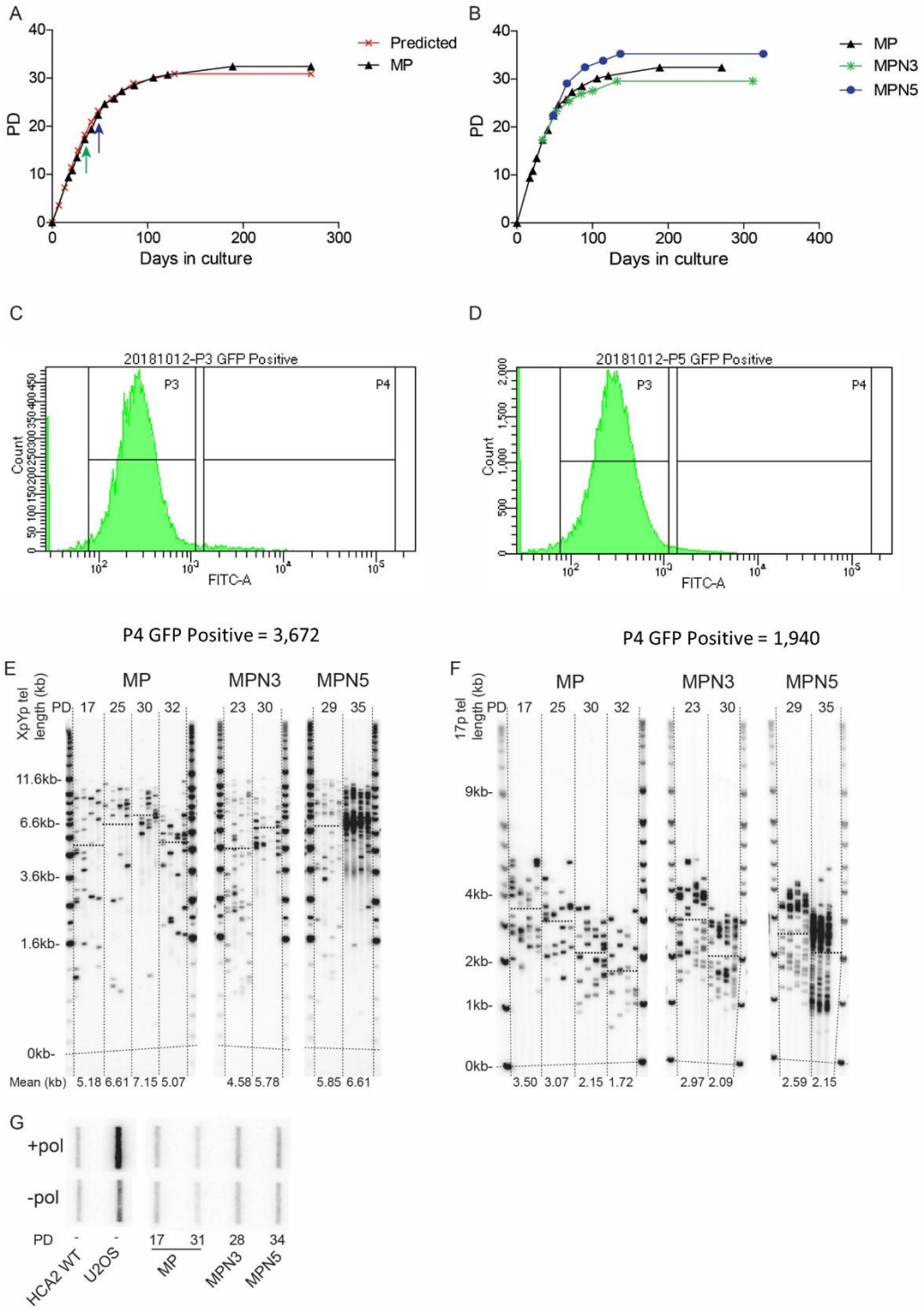


Figure 3.17: HCA2 WT mixed population cultures show no evidence of ATRX mutation and enter replicative senescence.

A) Growth curve of the expected curve (red) and the mixed population (black) for which ATRX KO was done at day 0 and with the 2 arrows indicating a second ATRX KO by CRISPR with the green arrow (PD17) and blue arrow (PD22). PDs depicted on Y-axis. B) Growth curves of the MPs with the original MP in black, the MP nucleofected at PD17 (MPN3) in green and the MP nucleofected at PD22 (MPN5) in blue; PDs depicted on Y-axis. Cell sorter results for MPN3 (C) and MPN5 (D) with P4 = GFP positive (cells being collected). STELA profiles at the XpYp (E) and 17p (F) chromosome ends for the MP, MPN3 and MPN5, with the average telomere length across the bottom also represented as a dotted line on the blots and PD across the top. G) '+' and '-' polymerase slot blot results of C-circle assay for MP, MPN3 and MPN4.

Despite no evidence of an ATRX mutation and a clear induction of replicative senescence, these cells displayed some different behaviours compared to the WT controls and the WT ATRX clones. They displayed a mixed population of cells with the majority of these presenting with the typical long and spindle-like morphology characteristic of young proliferating fibroblasts, as opposed to flat and large senescent cells (figure 3.18 A, comparative C) (Mocali et al. 2005; Maes et al. 2009). The presence of a precipitate, only observed in cells undergoing a telomere-driven crisis in our laboratory, was also observed in all three cultures suggesting the onset of crisis despite cells having intact cell cycle regulators such as TP53 and Rb (figure 3.18 B, comparative D). Fusion analysis was rendered difficult due to low amounts of DNA available, nonetheless, rare fusion events were detected suggesting once more the induction of a telomere crisis (red arrows figure 3.18 E). These cells did however stop growing showing no signs of ALT upregulation or replicative immortality.

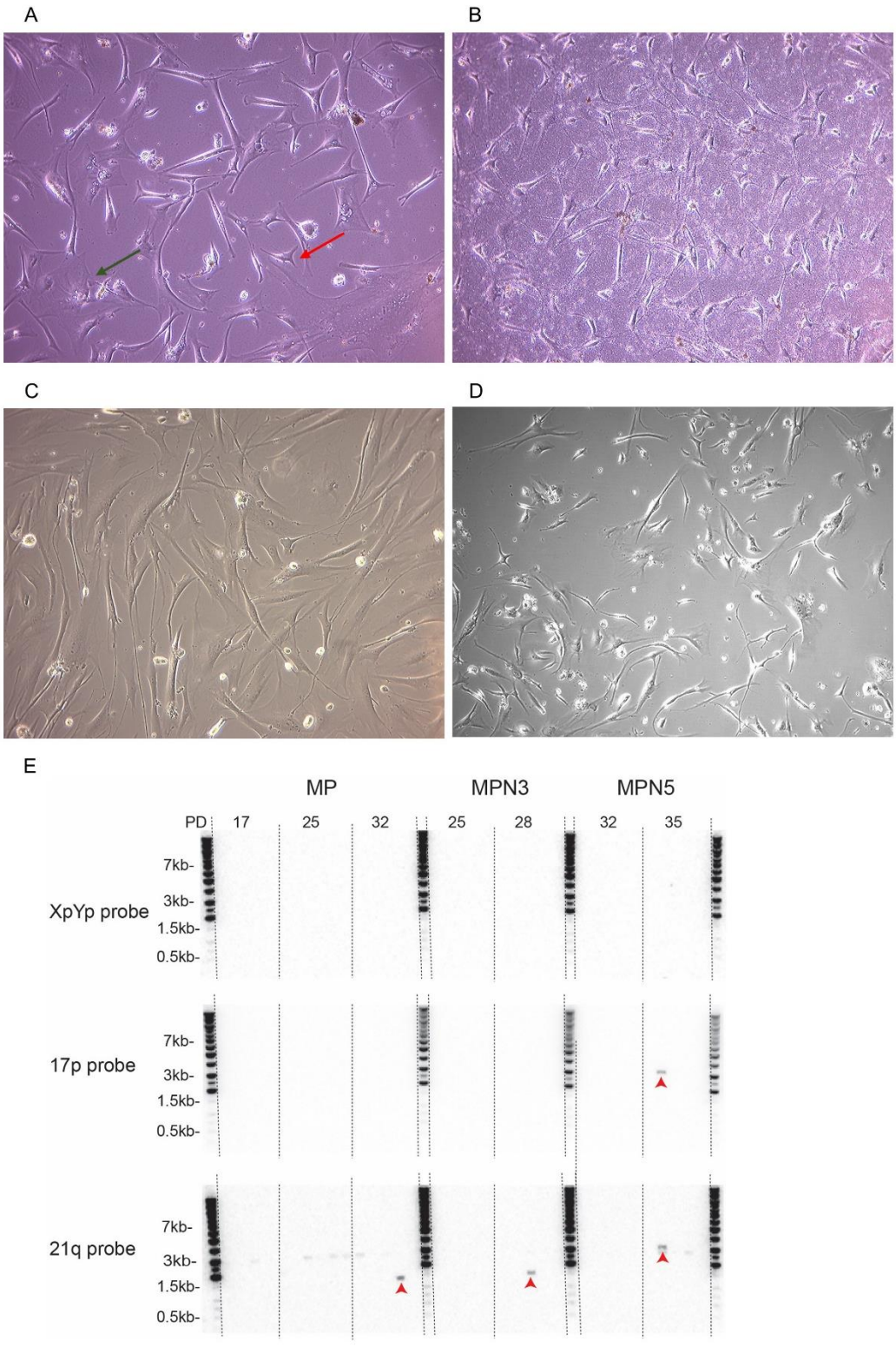


Figure 3.18: MP display characteristics of both senescence and crisis.
A) Picture at x10 magnification of the MP for which senescent cells were observed (green arrow) as well as long dividing cells (red arrow). B) Picture at x10 magnification of the MP showing heavy precipitate characteristic of crisis. C) Picture at x10 magnification of senescent HCA2 cells. D) Picture at x10 magnification of crisis HCA2 cells. E) XpYpM:17p6:21q1 fusion profiles of all MPs with fusions indicated by red arrows Hybridisation probes are detailed on the left. (Number of diploid genome equivalents analysed = $2x10^4$).

3.5 Discussion

3.5.1 Loss of ATRX and telomere stress is sufficient to upregulate and maintain ALT in fibroblast cells

The successful escape from a telomere-driven crisis is rare in fibroblasts and the data presented above, showed the immortalisation, following loss of ATRX, of primary human fibroblasts through the initiation of the ALT mechanism. This was confirmed by the presence of C-circles, in all clones for which ATRX was genetically lost, resulting in complete ablation of ATRX protein expression. Furthermore, upon long-term culture of escaping clones, all proliferated rapidly in the absence of telomerase, indicating a maintenance of the ALT mechanism for survival.

In contrast, WT cells, with functional TP53, showed a normal senescent phenotype despite the *ATRX* gene being similarly targeted. The nucleofection efficiency rate was dramatically reduced in the WT cells compared to the cells containing HPV E6E7 viral oncoprotein. A proportion of cells in these populations did however show some manifestations of crisis, including the presence of precipitate and telomere fusions, however, cells were unable to immortalise, presumably due to the ATRX protein remaining functional. TP53 has been found to be commonly mutated in cancers overall and has often been reported to be mutated, alongside ATRX, in ALT cancers suggesting that it may be required for ALT upregulation providing a further explanation for the successful escape from crisis in the HPV E6E7 clones (Hollstein et al. 1991; Voon et al. 2018; Oppel et al. 2019). This in turn also suggests that the loss of TP53 facilitates the successful targeting of ATRX by CRISPR/Cas9.

Furthermore, CRISPR targeted technology works by inducing strand breaks which in turn get repaired by host repair machinery inducing a mutation at the cut site. These have been associated with a high level of off-target effects, that limit the use of this technology in the clinic (Fu et al. 2013). These effects could explain the induction of crisis in the MP cultures despite ATRX remaining functional and ultimately leading to cell death rather than immortalisation. Repeating the knockout of ATRX earlier in the lifespan of these cells and establishing clonal populations with defined ATRX mutation will allow testing the role of ATRX in the induction of replicative senescence. In addition, due to the low efficiency of nucleofection in these cells, an alternative method to knock out or knock down ATRX could be considered. In example, the use of shRNA targeting by viral integration of the RNAi into the host genome resulting in its transcription, translation and export to the cytoplasm where it will bind to the RNA-induced silencing complex (RISC) which in turn will degrade the

complimentary ATRX coding mRNA has been successfully used in other studies (Moore et al. 2010; Napier et al. 2015). If cells reach crisis as a result, a nucleofection can be repeated to ensure a complete knock out of ATRX.

3.5.2 Residual ATRX expression inhibits the ALT mechanism

The data presented showed that despite generating clones with a mutated *ATRX* gene, these still produced a functional ATRX protein, although at a reduced level compared to the parental cell line. These clones were also unable to successfully immortalise compared to clones that had a complete loss of protein expression suggesting that ATRX is a strong inhibitor of the ALT mechanism. This agrees with other studies and supports the hypothesis that re-expressing ATRX in ALT clones would inhibit the mechanism of ALT leading to cell death (Clynes et al. 2015; Napier et al. 2015). In accordance with this, the HPV E6E7 control clones, which had a WT ATRX, simply underwent cell death due to a telomere-driven crisis in 100% of cases.

In addition, neither cells that had mutated or functional ATRX upregulated telomerase for survival, despite going through a telomere-driven crisis, which further supports the fact that fibroblasts preferentially upregulate ALT in most cases for survival and suggestive of a tight repression of telomerase activity in these cells (Lafferty-Whyte et al. 2009; Heaphy et al. 2011b). Furthermore, ALT positive cancers, mainly sarcomas originating from mesenchymal stem cells, have a mutated ATRX in the majority of cases (over 60%). A smaller subset undergo a DAXX mutation in conjunction or as a standalone mutation when ATRX is normal, mainly pancreatic neuroendocrine tumours (PanNETs), which have an even frequency of ATRX and DAXX mutations which are usually mutually exclusive (Heaphy et al. 2011a; Kim et al. 2017; Singhi et al. 2017). ATRX and DAXX work by chaperoning and depositing histone H3.3 at telomeres and loss of either protein results in instability due to an altered chromatin architecture, as well as an increase in replication stress due to the accumulation of secondary structures resulting in replication fork stalling and collapse (Episkopou et al. 2014; Wang et al. 2019). The loss of ATRX observed in ALT positive cancers therefore suggests that the lack of incorporation of histone H3.3 mediated by ATRX and DAXX is key in ALT upregulation (Heaphy et al. 2011a). In contrast, epithelial cancer cells rarely upregulate ALT (0.9%) and ATRX and DAXX statuses in carcinomas are not readily found in the literature (Heaphy et al. 2011b). Furthermore, the loss alone of ATRX in telomerase positive cells does not initiate ALT (Napier et al. 2015). This data shows that a substantially decreased level of ATRX expression

is enough to repress ALT implying that a low re-expression of the protein would be enough to inhibit this mechanism and therefore result in cell death (Haase et al. 2018; Yost et al. 2019). This is an interesting avenue for therapy as no switch to telomerase was observed implying that cells would be successfully targeted leading to cell death. It further suggests that expression of telomerase, more precisely the hTERT subunit, is tightly controlled in cells of mesenchymal origin hindering the upregulation of the enzyme for survival, despite cells undergoing a telomere-driven crisis and contributes to the maintenance of the ALT mechanism (Henson et al. 2002; Lafferty-Whyte et al. 2009).

3.5.3 Loss of ATRX affects cells before the onset of crisis

The loss of ATRX affects telomeres before the onset of crisis as seen with the increased heterogeneity at the first sampling points in the escapees compared to the controls. Despite these cells bypassing senescence, the first sample analysed for each clone corresponds to the PD point at which cells would normally senesce. This implies the presence of short telomeres triggering a cell cycle arrest and therefore suggests the presence of dysfunctional telomeres that may initiate the ALT mechanism generating the heterogeneity observed. To investigate whether the loss of ATRX or the approach to crisis are responsible for this heterogeneity, an earlier isolation of single cell clones would address this by providing more samples to analyse before telomeres become critically short. In addition, TP53 and Rb could be downregulated in a younger fibroblast population, enabling the knockout of ATRX to be carried out shortly after and therefore providing more sampling points before senescence and crisis.

In addition, as previously mentioned, ATRX is a chromatin remodeler which incorporates histone H3.3 at repetitive regions of the DNA and loss of it results in an altered chromatin structure especially at telomeres (Maze et al. 2013; Episkopou et al. 2014). Indeed, H3.3 incorporation at telomeres condenses and silences telomeric DNA thus preventing telomere transcription into TERRA. Therefore, the loss of ATRX leads to a more relaxed chromatin architecture thus exposing the ends of chromosomes to the DNA-damage response (DDR) as well as to transcription proteins (Goldberg et al. 2010). In addition, loss of ATRX results in replication fork stalling and collapse due to the accumulation of secondary structures, such as G4s and R loops, therefore rendering these telomeres more sensitive to replication stress. HDR-mediated DNA repair is commonly used to repair replication fork collapse which could explain the increased heterogeneity as telomeres are repaired through HDR (Costes and Lambert 2012; Clynes et al. 2014). This could also be evidence that telomere elongation in

the context of ALT, mediated through HDR, begins before crisis, as soon as telomeres become dysfunctional upon loss of ATRX (Wang et al. 2019).

3.5.4 Increased rate of fusion events in clones that escape crisis

As previously mentioned, escaping clones showed an increased rate of telomere fusions and this may arise due to more heterogeneous telomere length distributions with the presence of numerous short dysfunctional telomeres. The rate of fusions decreased after crisis despite telomeres remaining critically short and coinciding with the length at which fusions were previously seen. However, the lack of fusion events post-crisis was consistent with findings in the U2OS ALT positive cancer cell line, for which no fusion events were detected at all chromosome ends studied, despite the telomere length heterogeneity and the presence of telomeres within the length ranges at which fusions have been detected. These observations indicate that the initiation of the ALT phenotype is accompanied by a downregulation of NHEJ-mediated telomeric repair post-crisis (Koschmann et al. 2016).

The lack of fusion events in the context of ALT is surprising as ALT positive cell lines have been reported to be unstable throughout cancer progression and are associated with complex karyotypes, dysfunctional telomeres and microsatellite instability, compared to telomerase positive cells which stabilise telomeres after escape from crisis (Jeyapalan et al. 2005; Lovejoy et al. 2012; Jones et al. 2014; Li et al. 2019). Moreover, the presence of heterogeneous telomere length distributions, including short telomeres within the length range at which fusions occur, within ALT positive cells, such as HCA2$^{HPV\ E6E7\ ATRX-/-}$ escapees and U2OS cells, would be expected to be subjected to telomere fusion events. Nonetheless, ALT cells, that have undergone ATRX loss, silence NHEJ to favour HDR-mediated repair, which is consistent with the findings reported here, and suggest that in the absence of ATRX, short dysfunctional telomeres are recognised as DSBs and repaired through HDR exclusively rather than NHEJ thus preventing end-to-end fusions, although the mechanism of silencing is unclear (Koschmann et al. 2016). In accordance with this, in the absence of ATRX, the transcription of telomeres into TERRA is increased with TERRA migrating to short telomeres to induce HDR-mediated repair and elongation (Arora et al. 2014; Graf et al. 2017). This suggests that although the initial lack of ATRX does not appear to alter the fusion rate, perhaps the initiation of ALT, through the upregulation of C-circles and the successful escape from crisis, is the inhibitor of NHEJ to favour HDR as the use of the latter has been extensively reported

in relation to ALT in the literature (Dunham et al. 2000; Henson et al. 2002; Ait Saada et al. 2018; Zhang et al. 2019a).

The loss of ATRX in the context of ALT results in a decreased compaction of telomeric chromatin due to a reduced incorporation of histone variant H3.3 rendering telomeres more prone to transcription (Goldberg et al. 2010; Episkopou et al. 2014). Despite this, telomeres did not become dysfunctional at a different telomere length than the controls, as fusion events occurred when telomeres reached 1 kb and lower in all clones, despite telomeric chromatin being potentially more exposed to the DDR in ALT cells. This implies that the loss of ATRX is not directly responsible for the increased rate of fusion events in the escapees, but rather the resulting telomere length heterogeneity with a larger proportion short telomeres in the clones lacking ATRX.

3.5.5 Heterogeneity within escapees

Telomere length heterogeneity was observed in the escapees before and after crisis. There was also an increased heterogeneity amongst the escaping clones, with each clone presenting a different range of telomere lengths. Heterogeneity within telomere length in ALT positive cell lines has been reported previously and this data further supports this as a hallmark of ALT (Bryan et al. 1995). In addition, an increase in heterogeneity was observed when telomere lengths of all escaping clones post-crisis were combined, thus mimicking a polyclonal population, and compared to the ALT-positive cell line, U2OS, which showed no statistical difference at the 17p chromosome end. On this basis, a $HCA2^{HPV\ E6E7}$ mixed population could have complimented the above experiment through the study of telomere dynamics and to assess if the MP escaped crisis more readily than single cell clones.

The difference in heterogeneity between the chromosome ends studied revealed a difference in elongation as a result of ALT upregulation. It has been suggested that only short telomeres undergo telomere extension through ALT, however, the heterogeneity of the telomere length distributions before and after crisis and the lack of distinct allelic populations made it impossible to track the elongation events, and assess whether short telomeres are continually being elongated or whether telomere extension is a widespread phenomenon affecting all telomeres irrespective of telomere length (Teng et al. 2000). However, the 17p telomeres were more heterogeneous than the XpYp telomeres, suggesting the possibility that this chromosome end is subject to more extensive telomeric elongation events

compared to the XpYp telomere, and that the mechanism underlying ALT may be chromosome specific. These observations may also be consistent with the concept that a specific template is inserted in cis at each chromosome end during elongation of short telomeres and that inter-chromosomal exchanges are favoured for heterogeneity (Liu et al. 2018).

Heterogeneity in STING protein expression was also observed among all clones studied. Indeed, most escaping clones showed a downregulation or a loss of STING expression, however, one clone presented a similar intensity before and after crisis questioning the role of the cGAS-STING pathway in the ALT mechanism. Furthermore, STING appeared to be downregulated in clones that had residual ATRX activity and were not comparable to the parental protein expression further underlining inconsistencies in the involvement of this pathway in ALT. The involvement of the cGAS-STING pathway in the context of ALT has only been studied by one group and showed low residual activity in transformed fibroblast cells, similar to the cells presented above, and complete loss of STING in ALT positive cancer cell lines such as U2OS (Chen et al. 2017). This however does suggest that a downregulation of the pathway is enough to maintain the ALT mechanism although it is unclear why ALT negative clones have a downregulated STING expression. Further work investigating the upstream proteins of STING such as cGAS or the genetic sequence of the *STING* gene may clarify this issue. In addition, the level of cell death was not monitored post-crisis and the difference in levels of STING between clones could be indicative of multiple sub-populations within one clone, with some exhibiting intact STING and subsequent pathway and cell death. Analysis of the STING pathway may be informative upon longer time in culture, as U2OS cells for instance have undergone numerous PDs and is a well-established ALT-positive line, in comparison to these clones which have only undergone up to 18 PDs since crisis and upregulation of ALT.

3.5.6 Conclusions

In conclusion, the loss of ATRX combined with a telomere-driven crisis was sufficient to initiate and maintain ALT in human fibroblast cells in culture. Targeting ATRX in WT cells was unsuccessful and therefore establishing whether the loss of ATRX alone coupled with replicative telomere erosion was sufficient to initiate ALT was not achieved. However, the loss of ATRX affects cells before the onset of crisis in cells with abrogated TP53 and Rb function, implying that the loss alone would initiate a heterogeneity within telomere lengths,

which in turn may initiate crisis due to the accumulation of short dysfunctional telomeres. This also shows the importance of the ATRX protein in telomere stability and loss of this protein appears to occur early in the initiation of ALT. Finally, the heterogeneity in telomere length observed in escaping clones was consistent with an increased rate of fusion events as crisis was initiated, followed by a decrease in fusions by the last sampling points, consistent with observations in ALT positive cell lines. This suggests that fusion events detected are associated with crisis rather than the ALT mechanism and that ALT cells stabilise upon escape from crisis.

Chapter 4

Establishing the effects of the loss of ATRX on epithelial cancer cells and their ability to escape crisis

4.1 Abstract

Cancer cells gain replicative immortality through the upregulation of TMMs. The majority of tumour types (80-85%) utilise the telomerase mechanism that synthesises telomere repeats *de novo*. A further 10-15% of tumours activate the ALT mechanism. The *ATRX* gene has been strongly linked to the mechanism of ALT, as it is often mutated in cancers that upregulate this pathway.

The role that ATRX and telomere dysfunction play in the initiation of the ALT phenotype was examined. HCT116$^{ATRX-/-}$ cells were transfected with a DN-hTERT to abrogate telomerase activity and induce a telomere-driven crisis. A total of 149 single cell clones were isolated of which 33 escaped crisis and immortalised. Of those escapees, 4 immortalised using the ALT mechanism, with the remaining 29 upregulating telomerase. With ALT escapees, chromosome specific elongation of short telomeres ranging from 0.9 kb at XpYp to 6 kb at 17p was observed. ALT activity was further confirmed by the presence of C-circles, however, it was not maintained; the telomeric alleles that had been lengthened were subject to end-replication losses and gradual telomere erosion and, ultimately, telomerase activity was re-established in these clones. Furthermore, ALT positivity was observed in clones that died, through the C-circle assay, implying that ALT activity did not always lead to immortality, suggesting other underlying events involved in this pathway.

In conclusion, the genetic loss of ATRX combined with telomeric stress is sufficient to initiate the ALT mechanism temporarily in epithelial telomerase positive cancer cells, but this alone, is not sufficient to maintain telomere length and establish replicative immortality in the long-term.

4.2 Introduction

As previously described, cancer cells undergo a crisis phase prior to immortalisation through the activation of telomerase or the upregulation of the ALT mechanism. Epithelial cancer cells predominantly obtain replicative immortality by preferentially upregulating or reactivating telomerase activity. However a small percentage of carcinomas utilise the ALT mechanism, including hepatocellular carcinoma (7%), squamous cell carcinoma (2%) and a small subset of breast ductal carcinoma (Subhawong et al. 2009; Heaphy et al. 2011b). In contrast, there is no evidence of ALT upregulation in adenocarcinomas from multiple organs such as colon, lung, prostate, pancreas and cervix (Heaphy et al. 2011b). The function of telomerase has been well characterised, as this enzyme is active in stem cells and in the germline. The mechanism by which telomerase is upregulated in cancer cells has also been well established and can occur in multiple ways such as an amplification of the catalytic subunit of telomerase, *hTERT*, locus or activation through mutations of the hTERT promoter region that provide more binding sites for proteins including transcription factors which in turn initiate transcription of the gene, commonly observed in melanoma (Zhang et al. 2000; Huang et al. 2013).

In contrast, the mechanisms underlying the ALT pathway are less well understood and there is limited evidence of its activity in normal human cells (Berardinelli et al. 2010; Coluzzi et al. 2017; De Vitis et al. 2019). Mutations in the ATRX/DAXX/H3.3 complex have however been commonly associated with this mechanism (Heaphy et al. 2011a; Schwartzentruber et al. 2012). The ATRX protein binds to the DAXX protein which in turn binds to the histone H3.3 and ATRX chaperones the complex to repetitive regions in the genome, which include telomeres, to incorporate the H3.3 histone (Maze et al. 2013). This is considered to prevent the formation of G-quadruplex and R-loop structures and therefore limits replication fork stalling and collapsing at these fragile sites (Rhodes and Lipps 2015).

The incidence of ATRX mutations is most common in cells that upregulate ALT, and ATRX is therefore a strong candidate gene to investigate to obtain a better understanding of the mechanism by which ALT is induced in these cells (Lovejoy et al. 2012). Establishing how epithelial cells undergo a telomere-driven crisis and escape is important for understanding the cellular transition into malignant growth. Telomerase has been the focus of several novel targeted therapies as this could target 85% of cancers and telomerase is not expressed in normal somatic tissue (Kim et al. 1994). These include a wide range of approaches for which some have reached clinical trials such as suicide gene therapy through the delivery of

recombinant adenoviruses programmed to divide in cells expressing the catalytic subunit of telomerase only and therefore inducing apoptosis of the malignant cells (Song et al. 2003; Nemunaitis et al. 2010). Another approach is the binding of short oligonucleotides to the RNA binding site in the catalytic subunit of telomerase therefore preventing telomerase from extending telomeres. This in turn induces an erosion of the telomeres leading to cell death (Marian et al. 2010; Roth et al. 2010).

However, establishing whether epithelial cells can switch to the ALT mechanism following telomerase downregulation is important in the context of telomerase-targeted therapeutics. Furthermore, there are currently no ALT-targeted therapies and therefore gaining insight in ALT upregulation and maintenance could enable the development of diagnostic and prognostic tools, as well as, therapeutic targets for cancers that upregulate ALT (Pompili et al. 2017).

4.3 Aims

The hypothesis for this chapter is that the absence of ATRX activity combined with telomere stress is required to initiate the ALT mechanism. To address this, HCT116$^{ATRX-/-}$ cells, which are colorectal epithelial carcinoma cells that have upregulated telomerase to maintain their telomere length and gain replicative immortality, were used in all experiments. The focus of this chapter is to examine whether the induction of a telomere-driven crisis, by abrogating telomerase activity, in cells that lacked functional ATRX would affect the ability of cells to escape crisis and immortalise. Furthermore, specific chromosome ends were studied to track telomere length changes as ALT is initiated.

The aims for this chapter were to:

- Assess the effects of the ATRX knock out in single cell clones taken through a telomere-driven crisis following telomerase abrogation
- Establish what telomere maintenance mechanism was upregulated in clones that successfully escaped crisis
- Investigate telomere dynamics in ALT surviving clones to determine if ALT occurs at multiple chromosome ends and if it is chromosome specific
- Evaluate if clones that had previously escaped crisis by upregulating ALT can repeat this phenomenon if taken through another telomere-driven crisis
- Confirm if ALT can be maintained in surviving clones
- Consider the involvement of other mechanisms in the upregulation and maintenance of ALT such as the cGAS-STING pathway

4.4 Results

4.4.1 Analysis of HCT116$^{ATRX-/-}$ parental cell line and puromycin clones

HCT116 cells are colorectal carcinoma cancer cells from a male donor. These cells have active telomerase to maintain their telomere length. The *ATRX* gene was knocked out by exon replacement with rAAV by the Hendrickson laboratory (University of Minnesota) as previously described (Napier et al. 2015). The HCT116$^{ATRX-/-}$ parental cell line used for these experiments was tested for the ATRX protein by Western blot and showed no residual ATRX activity compared to the HCT116 WT cell line (figure 4.1 A). Cells were kept in culture for 26 days to study the growth dynamics: cells grew at a rate of 0.83 PD a day (figure 4.1 B).

To test whether a viral integration impacted the growth of cells, HCT116$^{ATRX-/-}$ cells were transfected with a puromycin selection plasmid as controls. 9 single cell clones were picked and kept in culture for 56 days. These clones were unaffected by the viral integration as they maintained 0.82 PD per day similar to that observed in the parental cell line and they maintained telomerase activity as expected (figure 4.1 B and C).

The STELA profiles at the 17p and XpYp chromosome ends showed similar heterogeneity in telomere length to the parental cell line, but more heterogeneous distributions than observed in clonal populations in the absence of telomerase (Baird et al. 2003). These observations are consistent with the activity of telomerase couple with telomere end-replication losses generating telomere length heterogeneity in clonal populations (figure 4.1 D and E). There was no statistical difference between each puro clone 4 time-points (P value = 0.21 and 0.48 for 17p and XpYp respectively following Kruskal-Wallis test) providing evidence of a maintenance of the telomeres as opposed to an erosion, consistent with telomerase activity and further confirms that the viral integration had no effect on the growth of these clones.

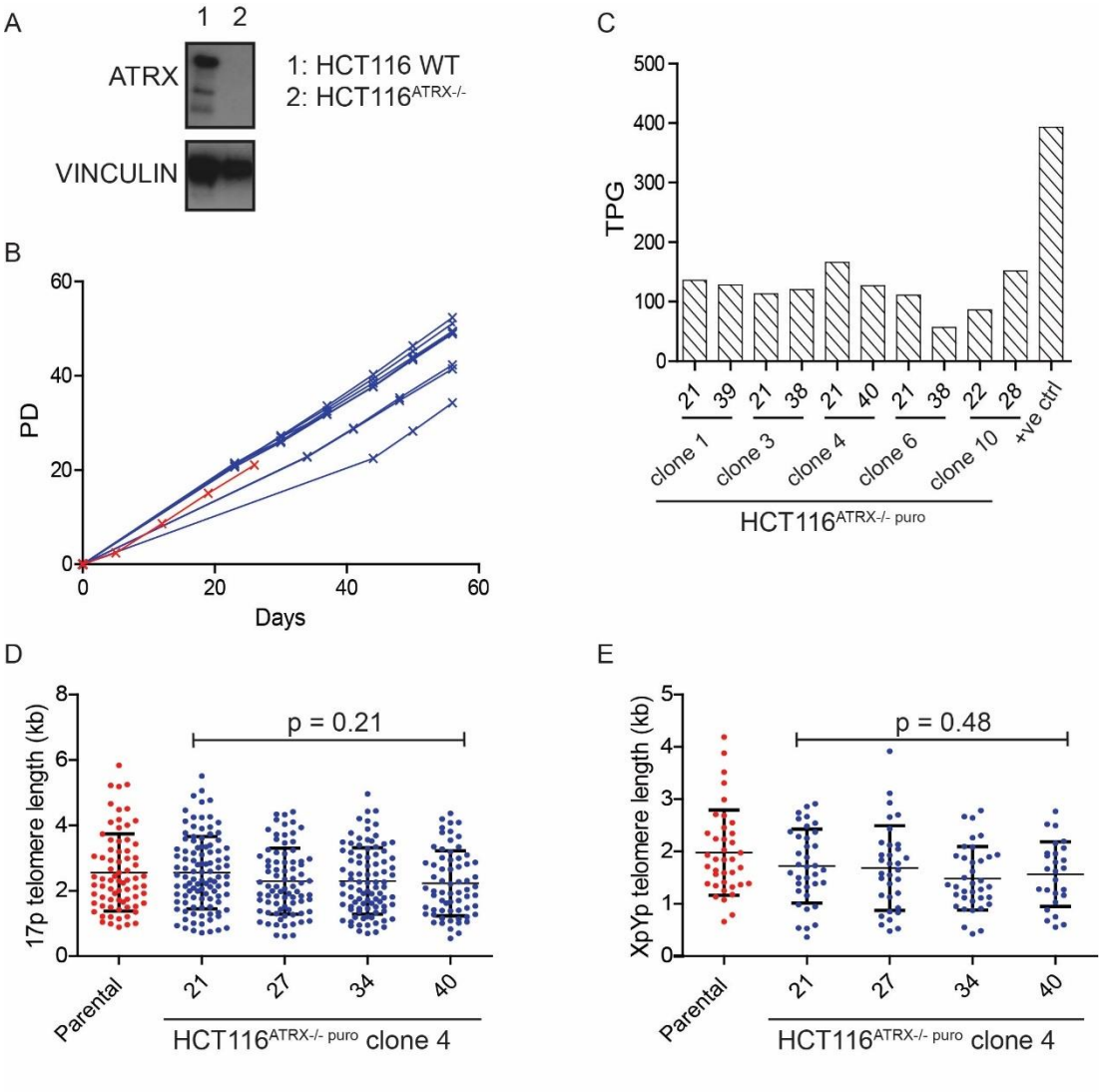


Figure 4.1: Characteristics of the HCT116$^{ATRX-/-}$ parental and puromycin control clones.
A) Western blot showing ATRX protein expression of the HCT116 WT and HCT116$^{ATRX-/-}$ parental cell lines and Vinculin protein expression as loading control. B) Growth curves of the HCT116$^{ATRX-/-}$ parental cell line in red and HCT116$^{ATRX-/-\ puro}$ clones in blue (n=9). PDs depicted on the Y-axis. C) Telomerase activity quantified by the TRAP assay showing the total product generated (TPG) in four of the puro clones at two time-points. D-E) Scatter plots of the length distributions at the 17p and XpYp telomeres with SD as error bars. The HCT116$^{ATRX-/-}$ parental cell line is highlighted in red and puro clone 4 at four different time-points (PD across the bottom) in blue. P value resulting from a Kruskal-Wallis test.

4.4.2 Induction of telomere stress in HCT116$^{ATRX-/-}$ cells

4.4.2.1 Effects of the DN-hTERT on cells

Prior work (undertaken by Mrs Julia Grimstead) had shown that upon transfection with a DN-hTERT to abrogate telomerase activity, HCT116$^{ATRX-/-}$ cells entered a telomere-driven crisis after a set period of normal cell growth (mean of 46 days and 26 PDs). Following transfection, cells were plated at limited dilutions (1:10; 1:100; 1:250; 1:500; 1:1,250) following addition of the DN-hTERT and 43 single cell clones were picked.

The DN-hTERT experiment was repeated three more times and a further 106 clones were obtained bringing the total to 149 single cell clones. From those, 38 went into crisis before any sample could be taken for cell counting and DNA extraction and are therefore not represented on the growth curves (figure 4.2). As the experiments were repeated, earlier samples were obtained between PD 16 and 19 to minimise the lack of representation of clones that underwent an early crisis.

The single cell clones entered a telomere-driven crisis, similar to that observed in HCT116 WT clones (Jones et al. 2014). Crisis occurred after a mean of 30 days of normal cell growth from the point of single cell cloning, representing a mean of 21 PDs. Crisis was determined by changes in morphology; from small dividing cells to large, multi-nucleated cells; and a slowdown in growth (figure 4.2).

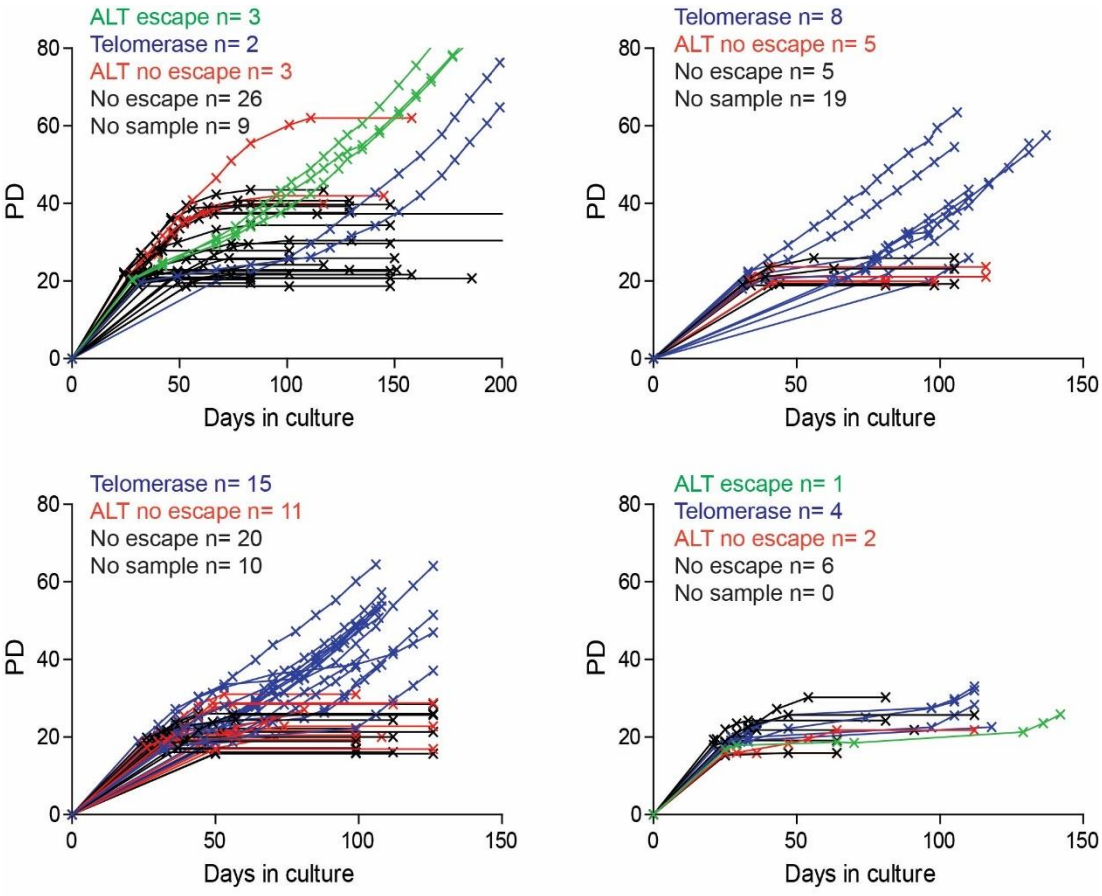


Figure 4.2: Growth curves of the 149 clones picked across four separate transfections.
The four infections are represented by the four separate panels. In green the ALT escapees (n=4); in blue the telomerase escapees (n=29); in red the clones that were C-circle positive and died (n=20); and in black the clones that were negative for C-circles and died (n=57). Additionally, there were 38 clones for which there were no samples and are therefore not represented on the graphs. PDs depicted on the Y-axis.

Expression of the DN-hTERT altered the telomere dynamics of the clones (figure 4.3), consistent with previous work (Jones et al. 2014). Prior to crisis, homogeneous bi-modal distributions were observed at the 17p chromosome end, representing bi-allelic telomere length distributions observed in single cell clones lacking telomerase activity (figure 4.3 A) (Jones et al. 2014). Some did however show overlapping allelic distributions exhibited through a heterogeneous distribution (clone 89 figure 4.3 A). This was followed by gradual telomere shortening, with a mean rate of erosion of 60 bps/PD between the first and the last samples before crisis across available clones (n=27). At the XpYp telomere, a single distribution was observed, consistent with the presence of a single detectable allele at this chromosome-end, and a mean rate of telomere erosion of 80 bps/PD was observed prior to the onset of crisis (n clones studied = 47) (figure 4.3 B). Overall, all clones underwent telomere erosion and a statistical difference was observed when comparing the mean

telomere lengths of all available clones at the first time-point with the last sampling point before crisis (p value of 0.03 and 0.0002 for 17p and XpYp chromosome ends respectively, following Mann-Whitney test) (figure 4.3 C).

Clones with long telomeres at the first sampling point underwent more PDs before crisis than those with short telomeres consistent with crisis being initiated by the accumulation of short telomeres resulting in a halt in cellular growth (Jones et al. 2014). Overall, long telomeres eroded more than short telomeres, as seen with clone 115 undergoing an erosion of 103 bps/PD at the XpYp chromosome end over 8 PDs, with a starting telomere length of 2.89 kb; and clone 117 only undergoing an erosion of 54 bps/PD over 9 PDs, with a starting telomere length of 1.31 kb (figure 4.3 B). In contrast, at the 17p chromosome end, allelic erosion was observed with some clones undergoing an equal erosion rate, clone 115 in example, whilst others showed differential allelic erosion, such as clone 117 with the longer allele eroding by 82 bps/PD and the shorter allele by 61 bps/PD. This can once more be associated with the starting telomere length as the long allele in clone 117 was 3.39 kb (mean) whilst the shorter allele was 1.46 kb (mean) further suggesting that long telomeres undergo more telomere attrition than short telomeres prior to crisis.

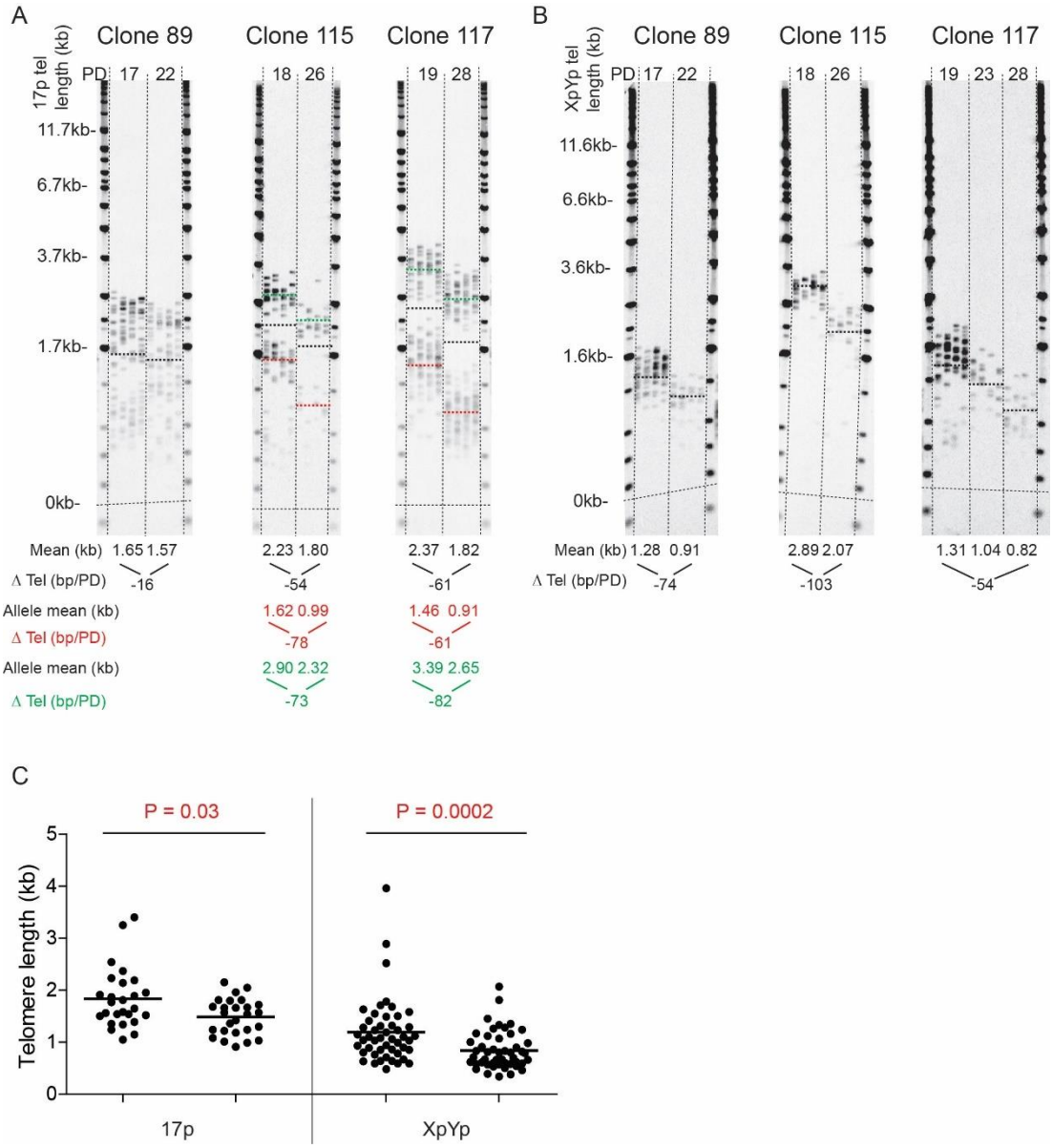


Figure 4.3: Telomere erosion following addition of the DN-hTERT.
STELA profiles at the 17p (A) and XpYp (B) chromosome ends in 3 single cell clones. Across the bottom, the overall mean telomere length in black, the mean telomere length of the shorter allele in red and the mean telomere length of the longer allele in green also represented as dotted lines on the blots. The overall rate of erosion as well as the allelic rate of erosion (where possible) are represented by the Δ Tel (bp/PD). PDs stated at the top. C) Scatter plot of the mean telomere lengths at the first available time-point for each clone and the last mean telomere length before crisis at the 17p and XpYp telomeres showing overall erosion across all clones. Statistical differences highlighted in red ($P < 0.05$)

4.4.2.2 Impact on survival rate

Previous observations have shown that all wild-type HCT116 clones expressing DN-hTERT, successfully escaped a telomere crisis following the re-activation of telomerase activity (Jones et al. 2014). In contrast, the absence of ATRX compromised the ability of the $HCT116^{ATRX-/-\ DN\text{-}hTERT}$ clones to escape crisis, with the majority (116 clones; 78%) of clones dying following the onset of crisis: of the 149 single cell clones picked, only 33 escaped crisis representing a 22% survival rate. Thus, the loss of ATRX affects the ability of cells to transit through a telomere-driven crisis and regain replicative immortality.

4.4.3 Analysis of escaping clones

4.4.3.1 Elongation at the XpYp and 17p chromosome ends

To investigate telomere dynamics of the clones that successfully escaped crisis, STELA profiles were generated at the 17p and XpYp chromosome ends for 29 escaping clones (figure 4.4 A and B). For 21 clones, pre- and post-crisis samples were available and analysed and 17 showed a statistically significant increase in telomere length heterogeneity at the XpYp chromosome end (p value = 0.03) when comparing before and after crisis standard deviations following an unpaired t-test (clones 92 and 111 figure 4.4 A and C). In contrast, no statistical difference was noted at the 17p chromosome end (p value = 0.7 following unpaired t-test) although the bi-allelic distributions observed before crisis were lost following immortalisation further consistent with an increase in heterogeneity in telomere length distributions (clones 92 and 111 figure 4.4 B and C).

In contrast, four clones (2, 3, 4 and 147) exhibited a telomere elongation to 1.98, 2.02, 1.66 and 1.69 kb respectively with a mean insertion of 1.06 kb (0.90, 1.21, 1.05 and 1.07 kb respectively) at the XpYp chromosome end with the entire distribution undergoing elongation (figure 4.4 A). At the 17p chromosome end, a bi-allelic distribution was observed, and the mean allelic telomere length was calculated alongside the overall mean telomere length (figure 4.4 B). The lower allele reached the critically short length of approximately 1 kb and appeared to undergo elongation to 6.99, 6.75 and 6.30 kb in clones 2, 3 and 4 respectively. Due to the lack of distinction between the two alleles in clones 2 and 3 prior to crisis and elongation, the overall telomere length was subtracted to the elongated allele mean length to establish the insertion length: an insertion of 5.64, 5.59 and 5.06 kb for clones 2, 3 and 4 respectively (mean of 5.43 kb). In addition, the longer allele prior to crisis simply

eroded as a function of cell division. In contrast, clone 147 exhibited an erosion of 17p telomeres with no elongation despite the observations at the XpYp telomeres. The elongation of short telomeres enabling further cellular division and survival in these clones does however suggest an ALT-like phenotype.

For the remaining 8 escapees, samples prior to the onset of crisis were not available due to the clones entering crisis early, nonetheless, the post-crisis STELA profiles exhibited a heterogeneous telomere length distribution similar to post-crisis samples of other escapees, and no statistical difference was noted when comparing post-crisis standard deviations (p value = 0.8 at both XpYp and 17p following an unpaired t-test) (figure 4.4 C).

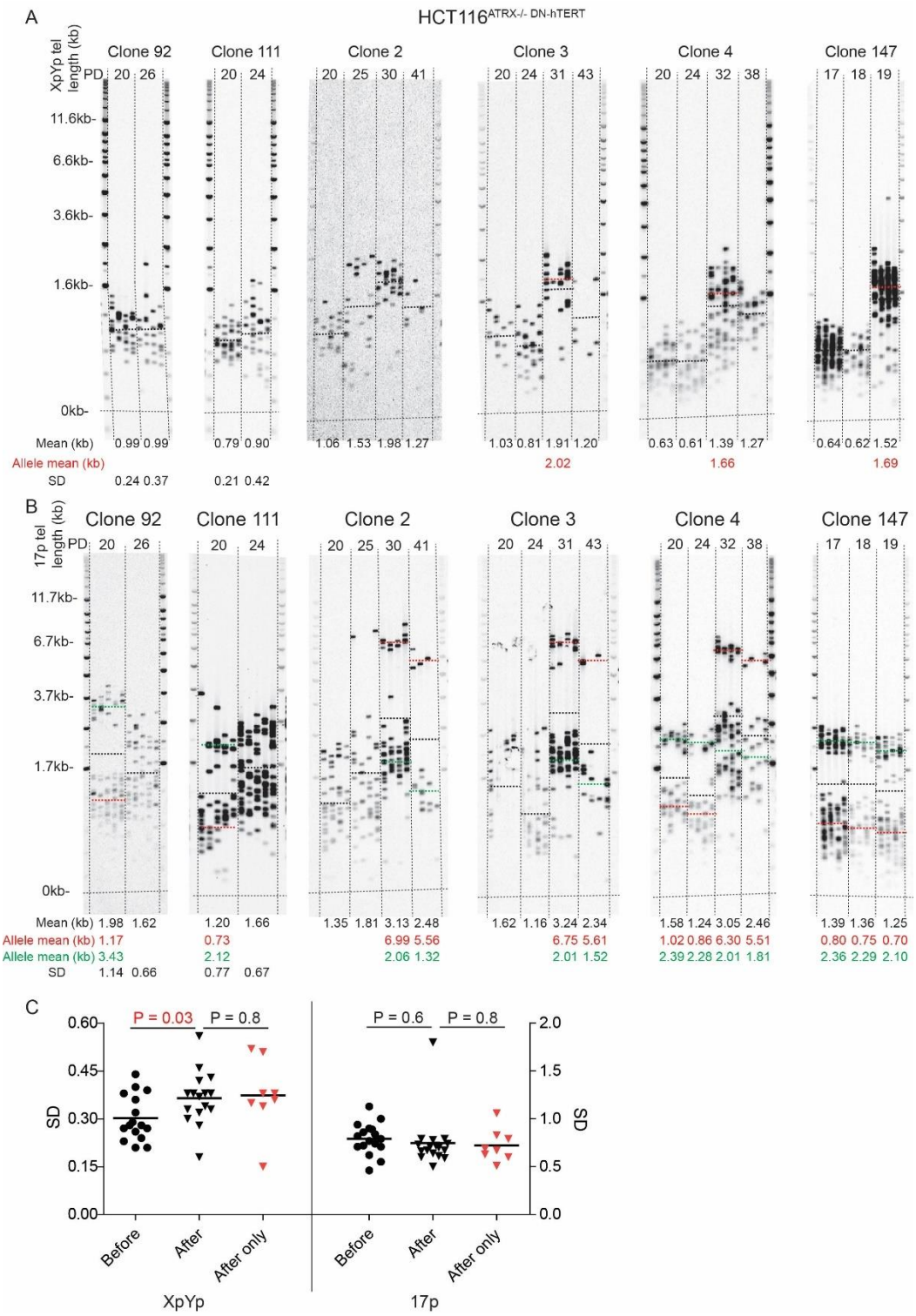


Figure 4.4: Telomere dynamics in escaping clones.
STELA profiles at the A) 17p and B) XpYp chromosome ends in 5 escaping clones. Clones 92 and 111 as examples of a change from homogeneous to heterogeneous telomere length distributions and clones 2, 3, 4 and 147 showing an ALT-like elongation. Across the bottom the overall mean telomere length in black; the shorter allelic mean telomere length prior to crisis that undergoes elongation in red; and the longer allelic mean telomere length prior to crisis in green, also represented as dotted lines on the blot. PDs stated at the top. C) Scatter plot of the standard deviations before (black circles) and after crisis (black triangle) and after crisis only for which no pre-crisis sample was available (red triangles) for the XpYp and 17p chromosome ends. Statistical differences are highlighted in red ($P < 0.05$).

4.4.3.2 Confirmation of ALT activity

To assess ALT activity in escapees, all surviving clones were subjected to the C-circle assay pre- and post-crisis. Clones 2, 3, 4 and 147, which exhibited telomere elongation at XpYp telomeres, were negative for C-circles before crisis however, a strong signal was detected at the time of elongation which confirmed ALT activity. Clones 2, 3 and 4 showed a gradual increase in C-circles after crisis with a maximal intensity, comparable to the positive control (U2OS), at PD 41, 43 and 31 respectively suggesting the formation of C-circles is gradual (figure 4.5 A). All results were normalised to the HCT116$^{ATRX-/-}$ parental cell line to establish positivity (figure 4.5 B).

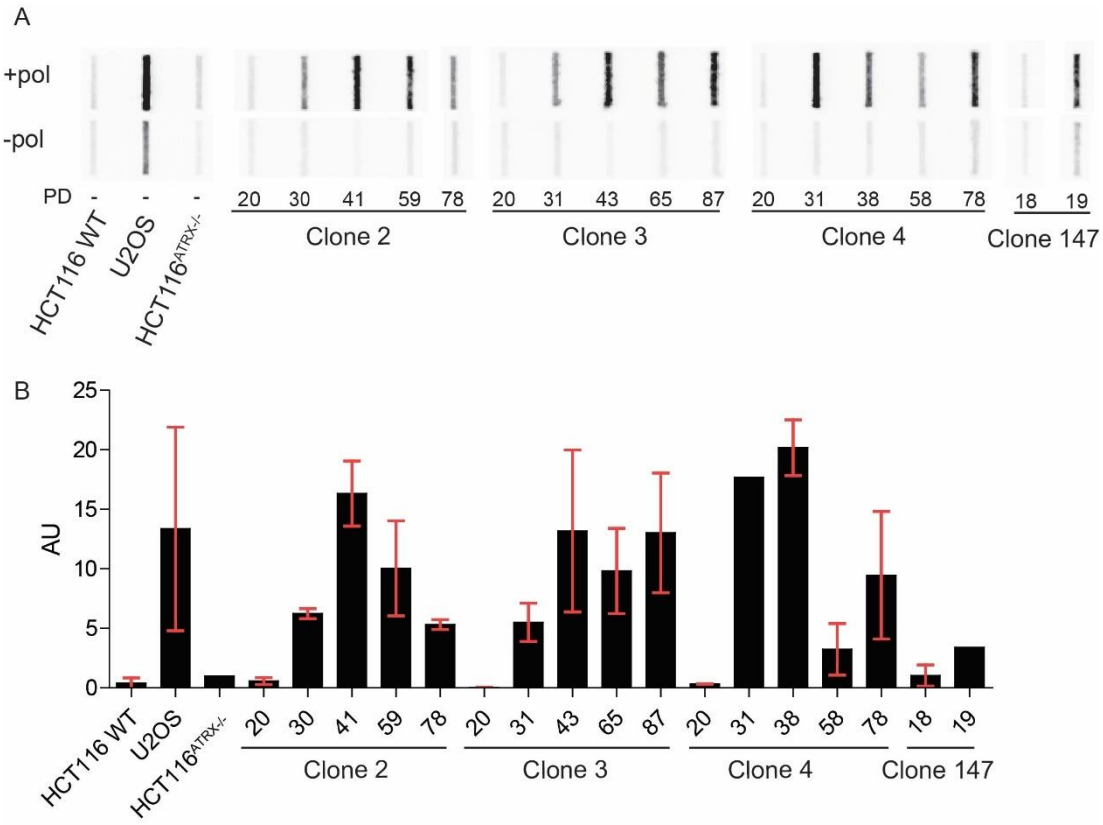


Figure 4.5: C-circle positivity in ALT-like clones.
A) '+' and '-' polymerase slot blot results of C-circle assay for clones 2, 3, 4 and 147, which underwent ALT-like elongation. B) Quantification of C-circle intensity normalised to the HCT116$^{ATRX-/-}$ parental and expressed in arbitrary unit (AU). SD as error bars where possible.

A further 3 clones (32, 48 and 131) were positive for C-circles (figure 4.6 A and B) despite the absence of an elongation event at the XpYp chromosome end (figure 4.6 C). For clones 32 and 131, a gradual increase in C-circles was observed as cells escaped crisis followed by a decrease by the last time-point suggesting C-circles were no longer being generated as telomerase was re-activated (figure 4.6 D). In contrast, clone 48 entered crisis before a sample was taken and the telomere dynamics before crisis were therefore unknown, however, this clone was strongly positive for C-circles post-crisis and this signal was maintained despite telomerase being active at PD 26 (figure 4.6 D).

The remaining 26 escapees were negative for C-circles and therefore ALT.

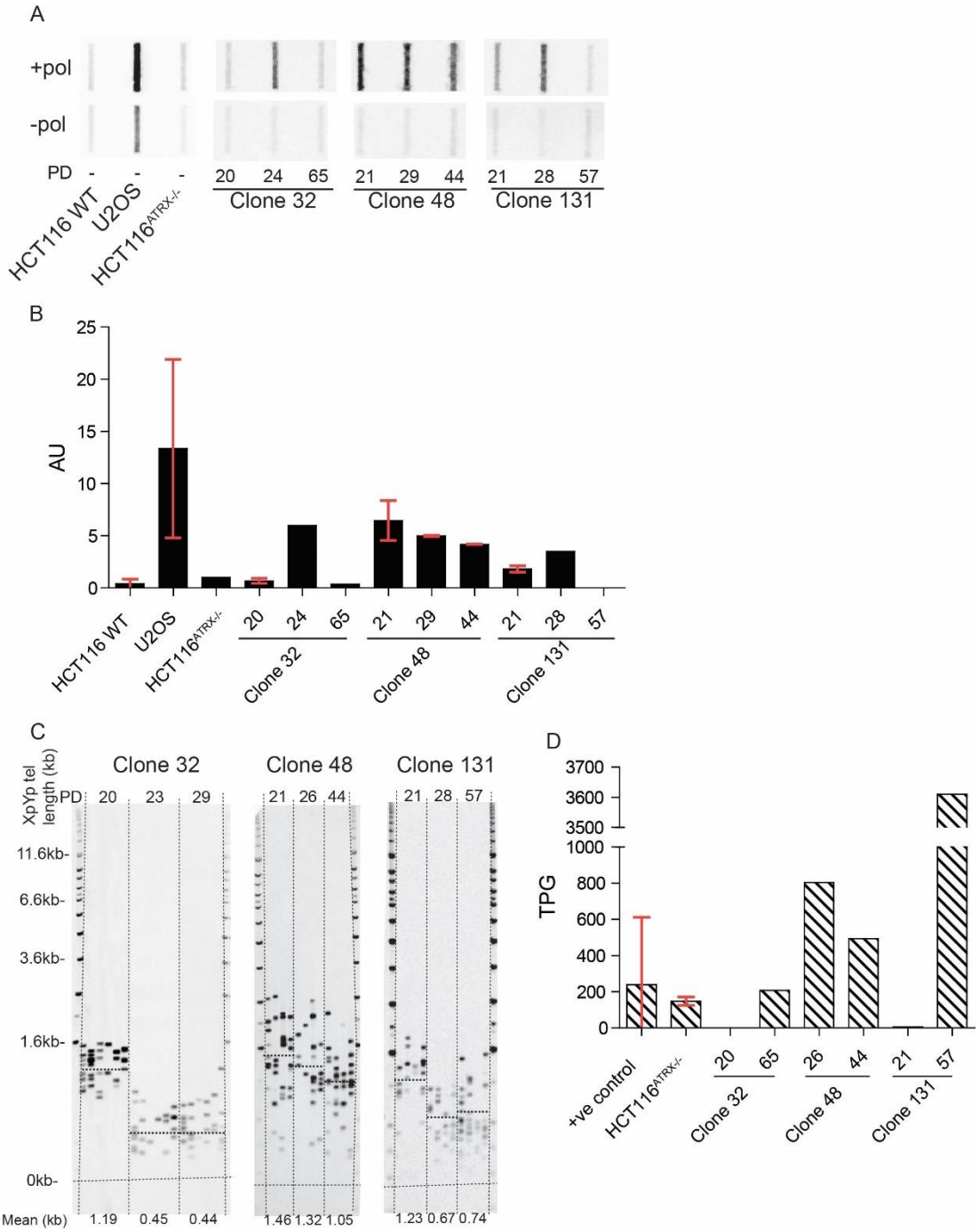


Figure 4.6: ALT-positive escaping clones with no ALT-like elongation.
A) '+' and '-' polymerase slot blot results of C-circle assay for clones 32, 48 and 131. B) Quantification of C-circle intensity normalised to the parental HCT116$^{ATRX-/-}$ expressed in arbitrary unit (AU). SD as error bars where possible. C) STELA profiles of clones 32, 48 and 131 at the XpYp chromosome end, in black the overall mean telomere length across the bottom, also represented as dotted lines on the blot. PDs stated at the top. D) Telomerase quantification expressed in total product generated (TPG). SD as error bars where possible.

4.4.3.3 Co-existence of telomerase and ALT activity

Telomerase activity was determined using the TRAPeze XL telomerase detection kit in all escapees. The 26 surviving clones that were negative for C-circles were strongly positive for telomerase activity after crisis (data not shown). There was no detectable telomerase activity in clones 2, 3 and 4 after crisis at the time of elongation, further suggesting ALT activity in these clones. A gradual increase in telomerase activity was however observed from PD 53, 58 and 51 respectively suggesting the upregulation of telomerase is ultimately required for long-term survival in these cells (figure 4.7 A-C).

Clones 2, 3 and 4 were kept in culture for 219 days for long-term analysis. These clones remained positive for C-circles at the last sampling point (PD 103, 114 and 103 respectively) despite telomerase being strongly re-activated by PD 83 (mean) thus suggesting a co-existence of both TMMs. In addition, clone 48 was strongly positive for C-circles as well as telomerase activity post-crisis. To establish if C-circle intensity was maintained in the long-term in the presence of telomerase activity, this clone was kept in culture for 228 days (figure 4.7 D). The results showed a maximal intensity of C-circles at PD 21 followed by a gradual decrease. Conversely, telomerase activity increased to reach maximal intensity at PD 44 with a return to levels comparable to the HCT116$^{ATRX-/-}$ parental by the last sampling point (PD 129) with a maintenance of the enzyme for survival. Unexpectedly, the intensity of C-circles increased at PD 67 and the signal was maintained until the last sampling point at PD 129 at which point cells were stored. These data show that both the telomerase and ALT mechanism can co-exist on the long-term which implies that the presence of C-circles does not inhibit telomerase activity, or vice versa.

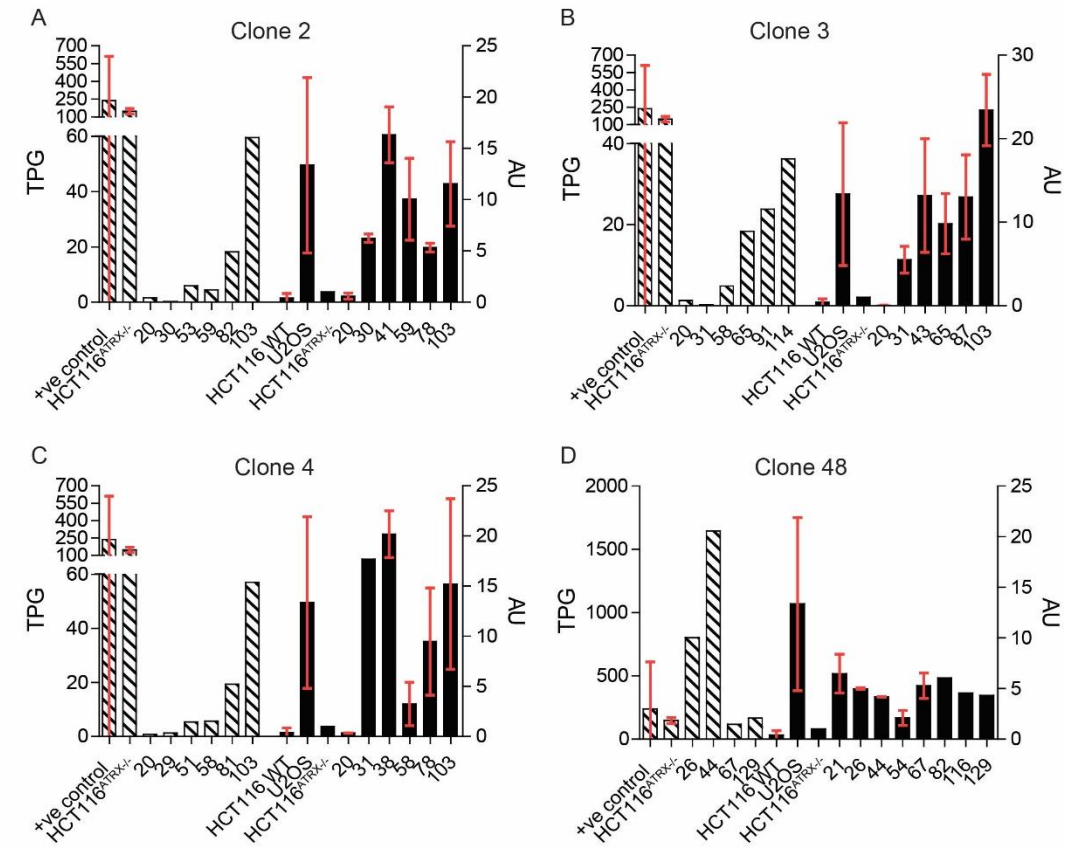


Figure 4.7: Co-existence of telomerase and ALT on the long-term in escaping clones.
Graphs showing a co-existence of telomerase and ALT through the levels of telomerase (left axis) measured in total product generated (TPG) and C-circle intensity normalised to the HCT116$^{ATRX-/-}$ parental (right axis) measured in arbitrary unit (AU) for HCT116$^{ATRX-/-\ DN\text{-}hTERT}$ clones 2 (A); 3 (B); 4 (C) and 48 (D). SD as error bars where possible.

4.4.4 Elongation events at multiple chromosome ends in ALT-like clones

4.4.4.1 Testing and optimisation

The elongation event observed at the 17p and XpYp chromosome ends provided evidence that ALT affects multiple telomeres. To investigate this further, STELA profiles were generated for an additional 9 chromosome ends (2p, 5p, 7q, 8q, 9p, 11q, 12q, 16p, 18q) to assess if ALT is chromosome specific in clones 2, 3 and 4. 12q and 18q showed no bands and very few bands in the control and these were therefore dismissed. 2p, 11q and 16p showed very few bands and were therefore also dismissed. For 5p, 7q, 8q and 9p, a PCR was repeated with an increased initial DNA concentration to maximise the result and establish the occurrence of an elongation event.

4.4.4.2 5p

The heterogeneous telomere length distributions indicate the presence of two allelic distributions at the 5p chromosome end, with similar mean telomere lengths. Telomeres were elongated to mean lengths of 3.87, 3.76 and 3.11 kb in clones 2, 3 and 4 respectively (an insertion of 1.89, 2.03 and 1.77 kb) (figure 4.8 and table 4.1).

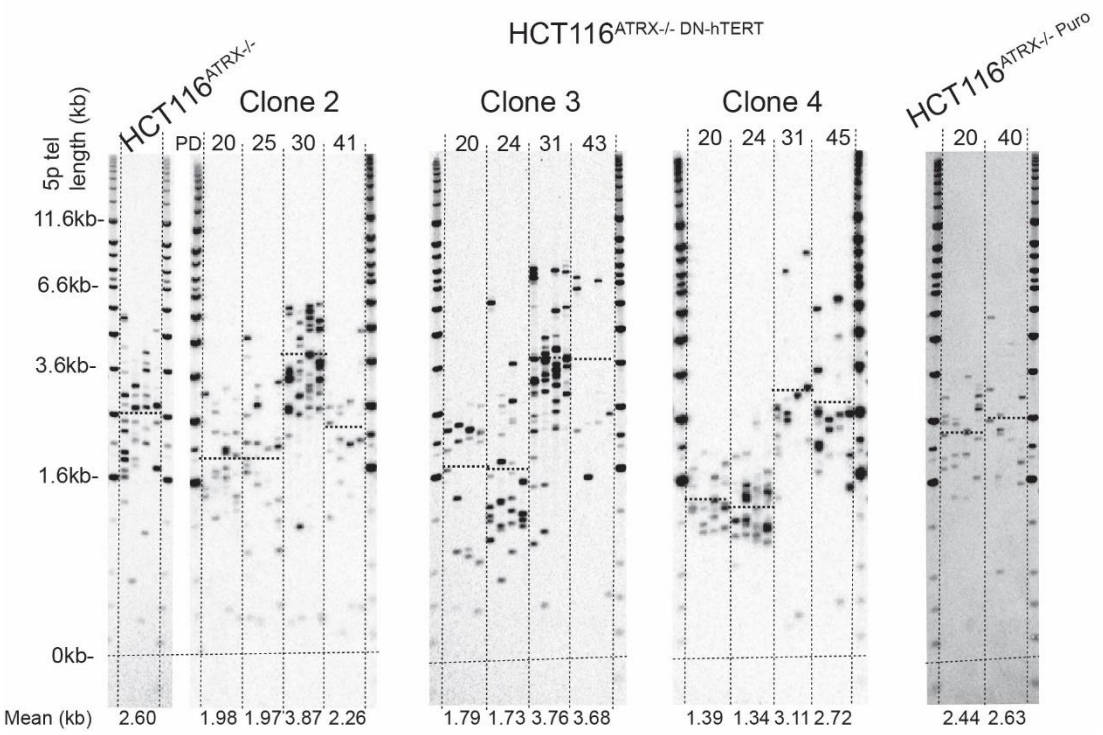


Figure 4.8: ALT-like elongation at the 5p chromosome end.
STELA blot showing 5p telomeres in the HCT116$^{ATRX-/-}$ parental; clones 2, 3 and 4; and in one HCT116$^{ATRX-/-\ puro}$ clone. In black the overall mean telomere length across the bottom also represented as dotted lines on the blot. PDs stated at the top.

4.4.4.3 7q

An initial DNA concentration of 250 pg per PCR reaction showed a mean of two bands per lane for each clone. The DNA concentration was increased to 1.25 ng per PCR reaction to increase the number of amplifiable molecules. This concentration was subsequently used for this primer for further analyses.

Similar to the 5p STELA profiles, heterogeneous telomere length distributions were observed at the 7q chromosome end, with no distinct allelic distribution again suggesting an overlap of two alleles. The 7q telomeres were elongated to 3.52, 3.32 and 2.81 kb in clones 2, 3 and 4 respectively (an elongation of 1.1, 1 and 1.32) (figure 4.9 and table 4.1).

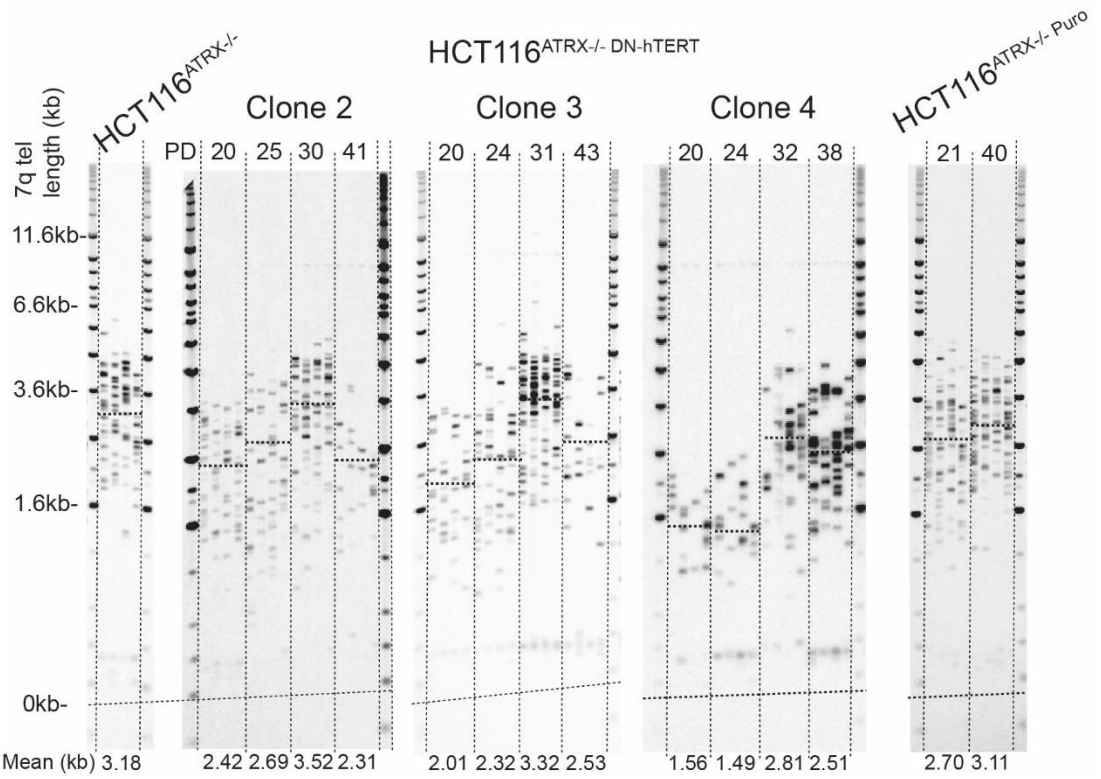


Figure 4.9: ALT-like elongation at the 7q chromosome end.
STELA blot showing 7q telomeres in the HCT116 $^{ATRX-/-}$ parental; clones 2, 3 and 4; and in one HCT116$^{ATRX-/-\ puro}$ clone. In black the mean telomere length across the bottom, also represented as dotted lines on the blot. PDs stated at the top.

4.4.4.4 8q

An initial DNA concentration of 250 pg per PCR reaction showed a mean of one band per two lanes for each clone. The DNA concentration was increased to 2.5 ng per PCR reaction to increase the number of amplifiable molecules. This concentration was subsequently used for this primer for other analyses.

The mean telomere length at the 8q chromosome end in the parental cell line was 7.25 kb. The mean telomere length in the three clones studied was 6.59 kb before crisis and 5.90 kb after crisis (figure 4.10). Thus, the 8q telomeres underwent erosion, as opposed to elongation, suggesting that the ALT mechanism only affects short telomeres.

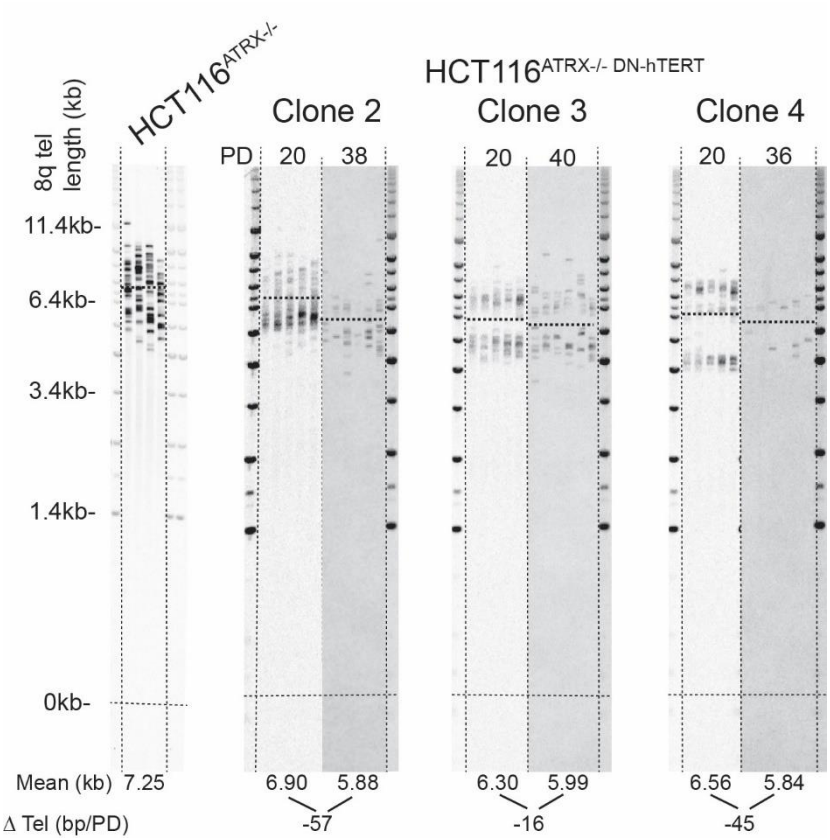


Figure 4.10: Erosion of long telomeres at the 8q chromosome end.
STELA blots showing 8q telomeres in the HCT116$^{ATRX-/-}$ parental; and in clones 2, 3 and 4 at two time-points (before and after crisis). In black the mean telomere length across the bottom, also represented as dotted lines on the blot, and the ΔTel showing the rate of erosion. PDs stated at the top.

4.4.4.5 9p

An initial DNA concentration of 250 pg per PCR reaction showed a mean of one band per two lanes for each clone. The DNA concentration was increased to 2.5 ng per PCR reaction to increase the number of amplifiable molecules. This concentration was subsequently used for this primer for other analyses.

The presence of short telomeres (below 0.7 kb) was observed in the parental cells, the puromycin clone and the ALT-like clones and these were therefore not associated to the ALT mechanism (highlighted in red box figure 4.11). As observed at the XpYp chromosome end, a single distribution was apparent at the 9p chromosome end and the entire distribution was elongated to 2.46, 2.21 and 1.43 kb in clones 2, 3 and 4 respectively (an insertion of 0.68, 0.68 and 0.59 kb) (figure 4.11 and table 4.1).

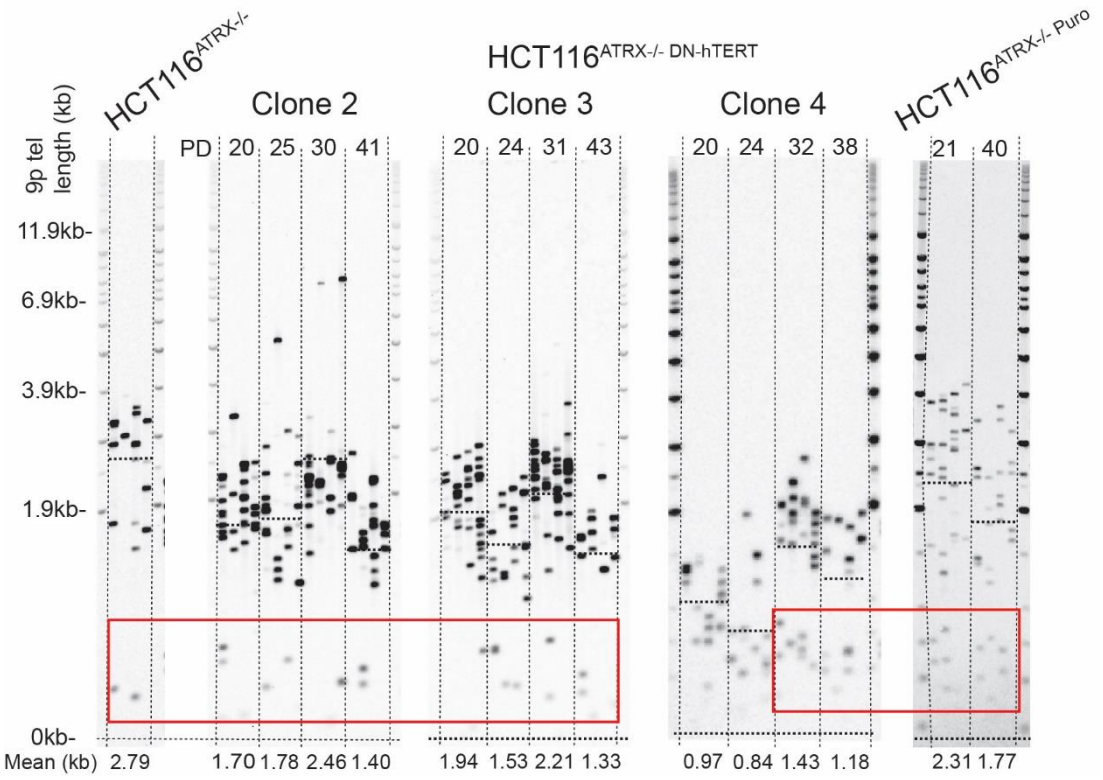


Figure 4.11: ALT-like elongation at the 9p chromosome end.
STELA blots showing 9p telomeres in the HCT116$^{ATRX-/-}$ parental; clones 2, 3 and 4; and in one HCT116$^{ATRX-/-}$ puro clone. The mean telomere length across the bottom; also represented as dotted lines on the blot. Short recurring telomeres in the red box. PDs stated at the top.

4.4.4.6 Chromosome specificity

All telomeres that underwent elongation did so at the same time: between PD 24 and 32; which implies that ALT affects all short telomeres at multiple chromosome ends. To assess

whether ALT-like elongation was chromosome specific, two values were compared for clones 2, 3 and 4 at the five chromosome ends that showed an elongation: the length at which the entire population or a single allele (17p) was elongated to post-crisis, and the length of the insertion summarised in table 4.1.

		Clone 2	Clone 3	Clone 4
Elongation length (kb)	5p	3.87	3.76	3.11
	7q	3.52	3.32	2.81
	9p	2.46	2.21	1.43
	17p (allele)	6.99	6.75	6.30
	XpYp	1.96	2.02	1.66
Insertion length (kb)	5p	1.89	2.03	1.77
	7q	1.1	1	1.32
	9p	0.76	0.68	0.59
	17p allele	5.64	5.59	5.06
	XpYp	0.90	1.21	1.05

Table 4.1: Consistent chromosome specific elongation in ALT clones.
Table showing the elongation length (kb) to which the single distribution or if a bi-allelic distribution was observed the single allele distribution was elongated to and the length of the insertion (kb) at each chromosome end studied in HCT116 $^{ATRX-/-\ DN\text{-}hTERT}$ clones 2, 3 and 4.

To calculate the insertion length, the shortest mean telomere length measured (PD 20, 24 and 24 for clones 2, 3 and 4) was subtracted from the longest mean telomere length measured (PD 30, 31 and 32 for clones 2, 3 and 4). Clones 2 and 3 were highly similar for both variables whereas clone 4 showed differences regarding the elongation length. The insertion length appeared to be consistent between each clone for specific chromosome ends.

4.4.5 ALT upregulation observed after taking same cells through second crisis

4.4.5.1 Elongation events seen at multiple chromosome ends

As shown above, chromosome specific telomeric elongation events were detected in several independent clones. In order to examine the reproducibility of these observations, the same clones were allowed to transit crisis for a second time. To do this, clones 2, 3 and 4 were taken back into culture from a freezing point prior to crisis (PD 23, 22 and 21 respectively) to establish if ALT upregulation could be repeated in clones that had previously escaped crisis

by means of ALT upregulation. These clones were renamed clones 2a, 3a and 4a. All clones were kept in culture for a varying time to allow cells to reach approximately 40 PDs (92 days for clone 2a, 77 days for clone 3a and 122 days for clone 4a) (figure 4.12 A). Clones remained in crisis for varying lengths of time: 30, 22 and 39 days for clones 2a, 3a and 4a respectively.

The three clones did however successfully escape crisis and elongation events were observed at all chromosome ends studied (5p, 7q, 9p, 17p and XpYp; figure 4.12 B-F). The STELA profiles were consistent with the previous observations despite mild variations due to samples being collected at different time-points and clones being therefore studied at different PD points. This provides further evidence that ALT elongation is chromosome specific.

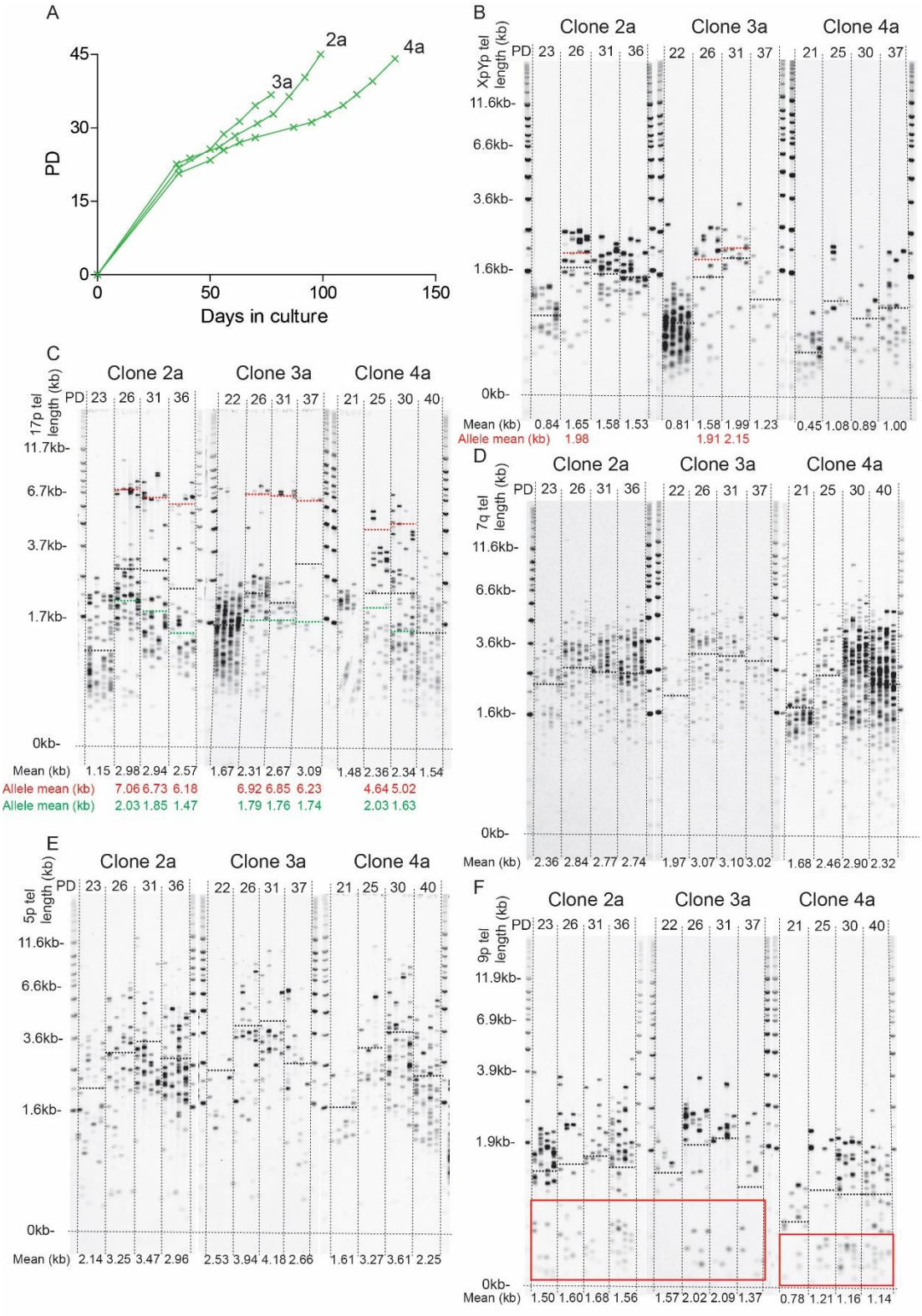


Figure 4.12: Characteristics of HCT116$^{ATRX-/-\ DN\text{-}hTERT}$ clones 2a, 3a and 4a taken through a second crisis. A) Growth curves of clones 2a, 3a and 4a with PDs depicted on Y-axis. STELA profiles at the B) XpYp; C) 17p; D) 7q; E) 5p; and F) 9p chromosome ends. In black, the mean telomere length; in red, the mean telomere lengths of the short allele before crisis that undergoes elongation; and in green the mean telomere length of the long allele before crisis that erodes across the bottom, also represented as dotted lines on the blot. Short recurring telomeres in the red box. PDs stated at the top.

4.4.5.2 ALT positivity and telomerase activity

To further confirm ALT upregulation in clones 2a, 3a and 4a, all clones were submitted to the C-circle assay. The formation of C-circles resembled that observed in clones 2, 3 and 4 with a gradual increase in signal post-crisis with a maximal intensity at PD 33, 31 and 30 for clones 2a, 3a and 4a respectively (figure 4.13 A). Telomerase activity before crisis was null followed by a sharp increase by the last sampling point further suggesting these cells require telomerase upregulation to survive on the long-term (figure 4.13 B).

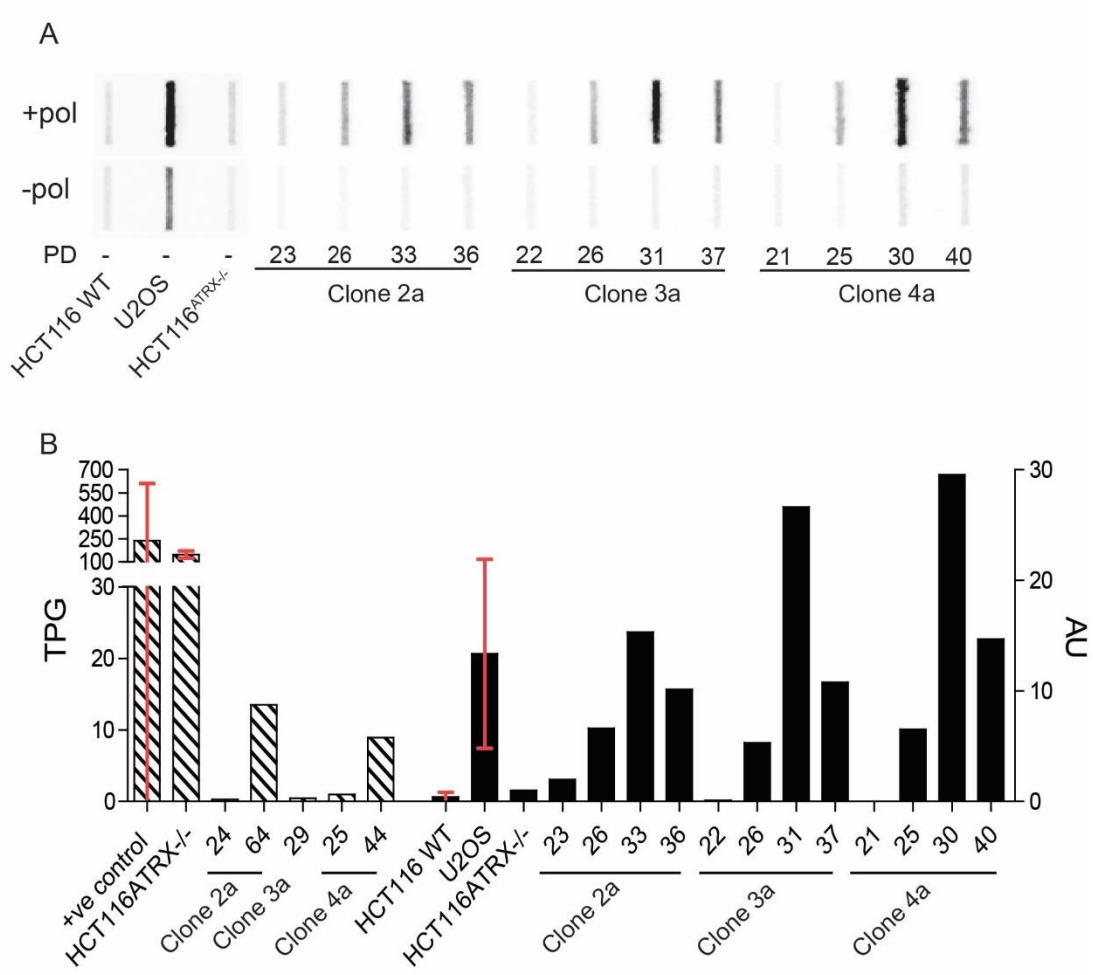


Figure 4.13: Transient ALT upregulation in HCT116$^{ATRX-/-\ DN-hTERT}$ clones 2a, 3a and 4a.
A) '+' and '-' polymerase slot blot results of C-circle assay for clones 2a, 3a and 4a. B) Graph showing the quantification of the total product generated (TPG) indicative of telomerase activity on the left axis and the C-circle quantification normalised to the HCT116$^{ATRX-/-}$ parental on the right axis expressed in arbitrary unit (AU). SD as error bars where possible.

4.4.5.3 Chromosome specificity

To further address the chromosome specificity of an ALT-like elongation, clones 2a (PD31), 3a (PD31) and 4a (PD30) were compared to clones 2 (PD30), 3 (PD31) and 4 (PD31) respectively to establish if the length at which telomeres were elongated to alongside the insertion length were consistent between clones that underwent an ALT-like elongation on two separate occasions. Elongated telomere lengths were plotted for each chromosome end for all clones (figure 4.14 A-C), and Mann-Whitney tests were carried out to assess any statistical differences between clones (figure 4.14 and summarised in table 4.2). Clone 3 showed no statistical differences with clone 3a upon comparison of elongated telomere lengths at all chromosome ends studied showing a consistent elongation event. Statistical differences were noted at XpYp, 7q and 9p for clones 2 and 2a and XpYp, 17p and 9p for clones 4 and 4a, whilst the remaining chromosome ends showed no differences in elongation. In addition, 5p telomeres showed no differences across all clones. Alongside, the insertion lengths were also plotted and clones 2, 3 and 4 were grouped and compared to clones 2a, 3a and 4a and no statistical difference was consistently observed, further supporting the hypothesis that ALT elongation is chromosome specific.

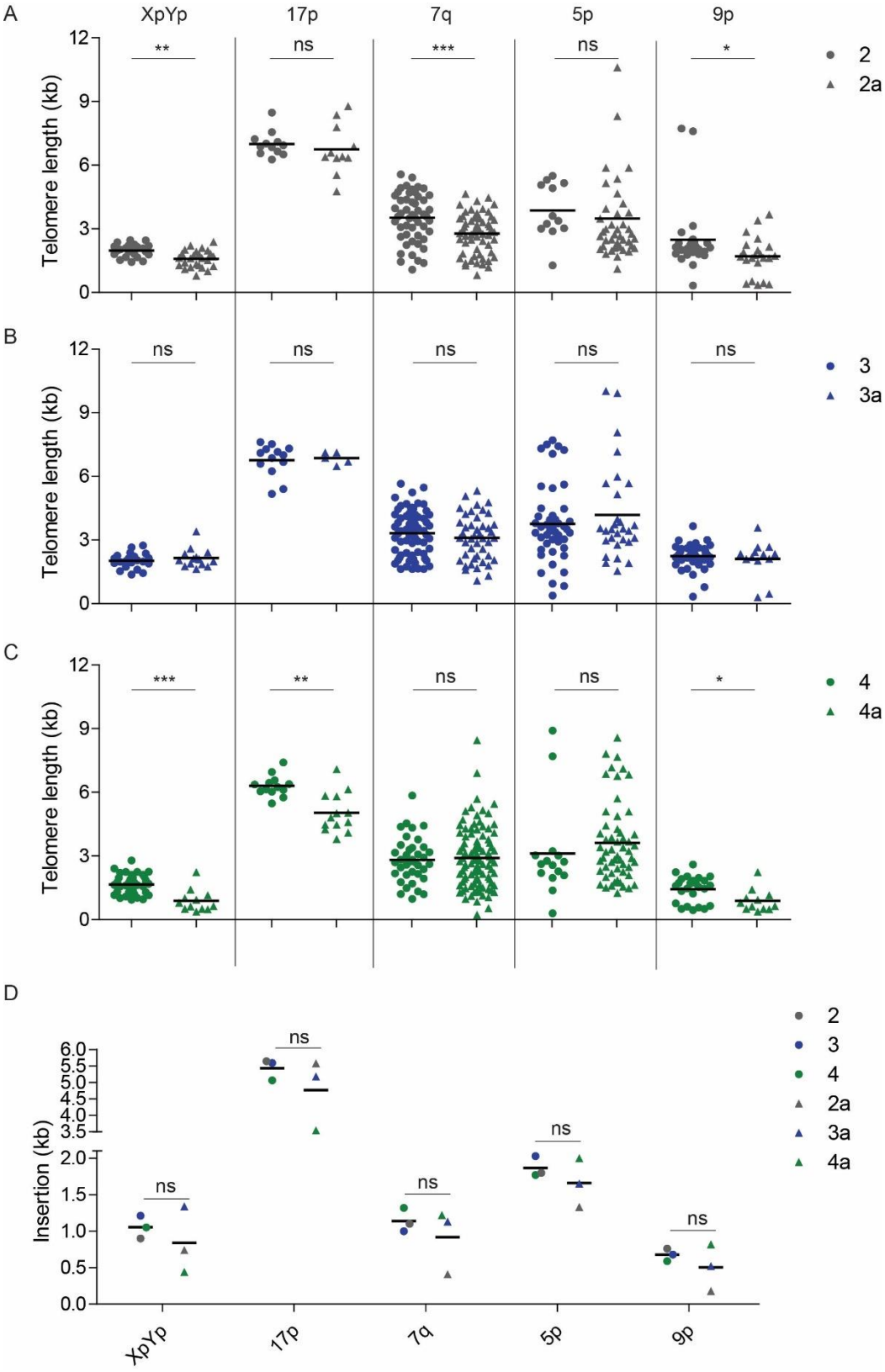


Figure 4.14: Chromosome specific ALT elongation.
Scatter plots showing the elongated telomeres for (A) clone 2 (grey circles) and clone 2a (grey triangles); (B) clone 3 (blue circle) and clone 3a (blue triangles); (C) clone 4 (green circles) and clone 4a (green triangles); at XpYp, 17p, 7q, 5p and 9p chromosome ends with the statistical difference resulting from Mann-Whitney test stated above. D) Scatter plot of the insertion length at the XpYp, 17p, 7q, 5p and 9p chromosome ends and the statistical difference stated above resulting from an unpaired t-test.

		XpYp	17p	7q	5p	9p
Elongated telomere length	2 vs 2a	0.0011	0.25	0.0005	0.054	0.024
	3 vs 3a	0.63	0.84	0.30	0.76	0.74
	4 vs 4a	0.0001	0.0015	0.97	0.19	0.013
Insertion	2, 3, 4 vs 2a, 3a, 4a	0.41	0.59	0.33	0.35	0.42

Table 4.2: Chromosome specific ALT elongation.
Table stating the p values comparing 2 and 2a; 3 and 3a; 4 and 4a elongated telomere lengths resulting from a Mann-Whitney test with p values < 0.05 in red. P value comparing 2, 3 and 4 with 2a, 3a and 4a insertion lengths resulting from an unpaired t-test.

4.4.6 Analysis of clones that did not survive

4.4.6.1 ALT-like elongation events in clones that did not survive

Clones 108 and 132 showed an elongation at the XpYp chromosome end to 1.88 and 1.52 kb respectively (an insertion of 0.69 and 0.78 kb) (Figure 4.15 A). At the 17p chromosome end, a bi-allelic distribution was observed at the first sampling point followed by the loss of an allele by the second time-point. Two long telomeric molecules were measured at 3.88 and 5.89 kb in clone 108 at PD 22 suggesting the initiation of an elongation, however, these disappeared by PD 25 (figure 4.15 B). These clones were positive for C-circles and negative for telomerase further indicating the initiation of the ALT mechanism however, these clones died despite observing ALT-like characteristics (figure 4.15 C and D). Signs of crisis such as the presence of large multi-nucleated cells appeared early in the culture of these cells resulting in a slow-down in growth. This was followed by a rapid cell growth following elongation of the XpYp telomeres indicated by the green box on the growth curves (figure 4.15 E). Furthermore, telomeres were elongated to a similar length across all clones (2, 2a, 3, 3a, 4, 4a, 108, 132 and 147) that underwent an ALT-like elongation at the XpYp chromosome end, apart from clone 4a, for which the mean telomere length was shorter. Furthermore, a statistical difference was observed between clone 4a and all other clones apart from 2a and 132 as a result of a Kruskal-Wallis test followed by a Dunn's multiple comparison test (summarised in table 4.3).

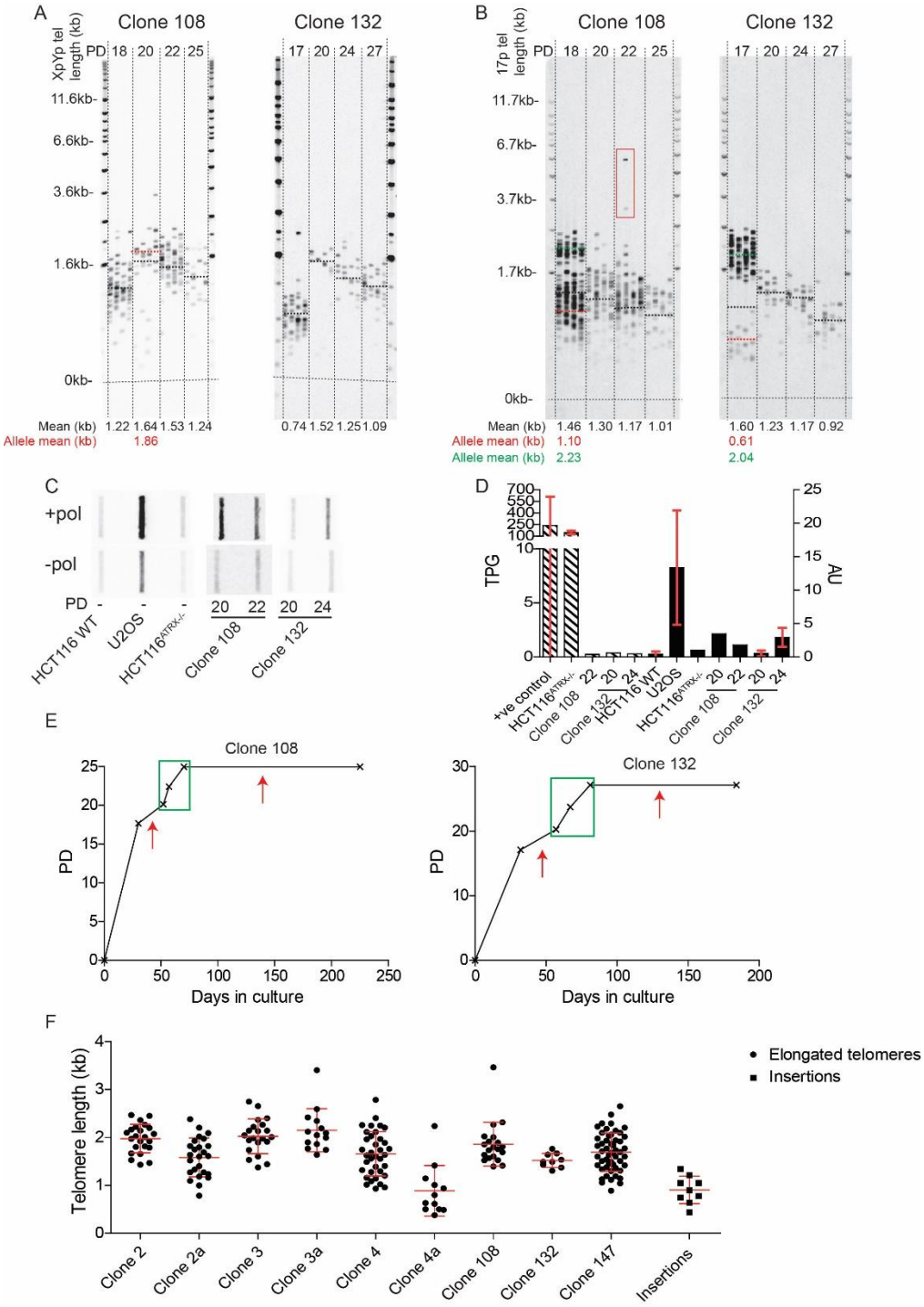


Figure 4.15: ALT-like elongation and C-circles in clones that died.
STELA profiles at the XpYp (A) and 17p (B) chromosome ends. In black, the mean telomere length; in red, the mean telomere lengths of the short allele before crisis that underwent elongation; and in green, the mean telomere length of the long allele before crisis that eroded, also represented as dotted lines on the blot. PDs stated across the top. C) '+' and '-' polymerase slot blot results of C-circle assay for clones 108 and 130. D) Graph showing the quantification of the total product generated (TPG) indicative of telomerase activity on the left axis and the C-circle quantification normalised to the HCT116$^{ATRX-/-}$ parental on the right axis expressed in arbitrary unit (AU). SD as error bars where possible. E) Growth curves of clones 108 and 132 with red arrows indicating periods of crisis and the green box a period of growth. PDs depicted on Y-axis. F) Scatter plot of the XpYp telomere lengths of all clones that exhibited ALT-like telomere elongation (circles) and the insertion length (squares). SD as error bars.

All	***								
	2	**2a**	**3**	**3a**	**4**	**4a**	**108**	**132**	**147**
2		*	ns	Ns	ns	***	Ns	ns	ns
2a			*	*	ns	ns	Ns	ns	ns
3				Ns	ns	***	Ns	ns	ns
3a					*	***	Ns	*	ns
4						*	Ns	ns	ns
4a							**	ns	**
108								ns	ns
132									ns

Table 4.3: Comparison of elongated telomere lengths at the XpYp chromosome end.
Table summarising the result of a Kruskal-Wallis test (all) followed by a Dunn's multiple comparison test comparing each pair of clones.

All remaining clones that died were therefore analysed through the C-circle assay to assess if ALT initiation was a more wide-spread phenomenon.

4.4.6.2 Analysis of the remaining clones

The C-circle assay was used across all clones for which samples were available (n=76). The intensity of the C-circle band is directly correlated to the amount of C-circles being generated and is therefore used as a quantifiable measure of ALT activity (Henson et al. 2009). The results showed a variation in intensity suggesting some clones had stronger ALT activity (figure 4.16 A). The XpYp telomere length distributions were determined in clones that were positive for C-circles (n=19) to establish the occurrence of an elongation event further confirming ALT initiation (figure 4.16 B). No obvious elongation was noted, however, clones exhibited different telomere length profiles. Indeed, a subset of clones (n=8), such as clone 130, showed a gradual increase in C-circle positivity which was accompanied by a period of sustained cellular growth (total of 28 PDs) (figure 4.16 A and C). STELA profiles revealed a starting mean telomere length of 1.49 kb in clone 130 suggesting the possibility of an elongation prior to the first sampling point (figure 4.16 B).

In addition, a small number of clones (n=3) were strongly positive for C-circles despite dying shortly after the first sampling point. For example, clone 52 only underwent 21 PDs before crisis, nonetheless, this clone had long XpYp telomeres (mean 1.59 kb). In contrast, 8 clones, such as clone 83, had short telomeres (mean 0.50 kb at last sampling point) and the C-circle intensity was fainter than the other C-circle positive clones suggesting that elongation of

short telomeres is required for sustained cellular growth and appears to be one the first steps in ALT initiation and directly correlates with C-circle intensity.

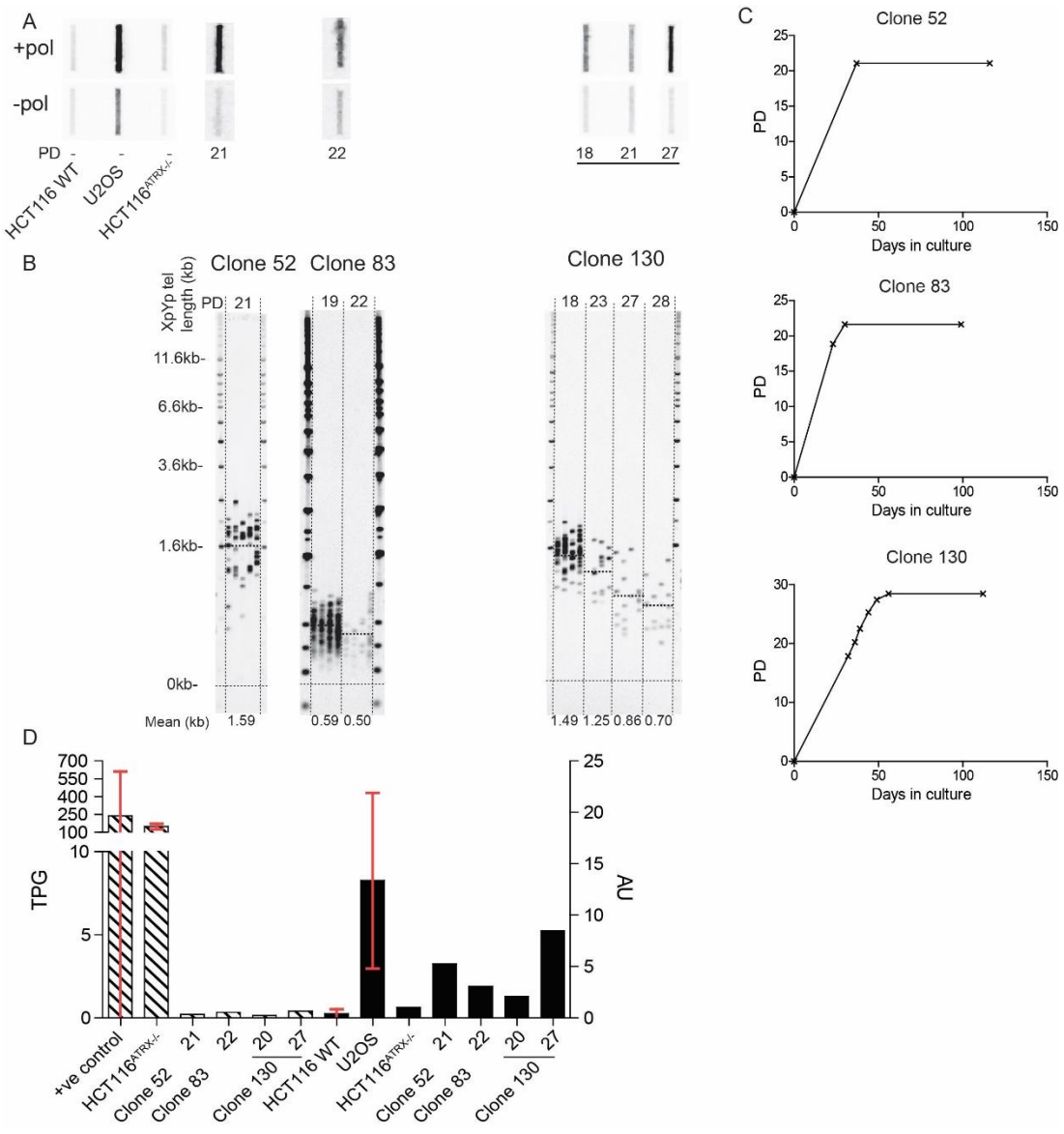


Figure 4.16: Characteristics of HCT116$^{ATRX-/-\ DN-hTERT}$ ALT positive clones that died.
A) '+' and '-' polymerase slot blot results of C-circle assay for clones 52, 83 and 130 and B) matching STELA profiles at the XpYp chromosome end, with the mean telomere length stated across the bottom also represented as dotted black lines on the blots. PDs stated at the top. C) Growth curves of clones 52, 83 and 130. PDs depicted on the Y-axis. D) Graph showing the quantification of the total product generated (TPG) indicative of telomerase activity on the left axis and the C-circle quantification normalised to the HCT116$^{ATRX-/-}$ parental on the right axis expressed in arbitrary unit (AU). SD as error bars where possible.

4.4.7 HCT116$^{ATRX-/-\ DN\text{-}hTERT}$ mixed population

Once all single cell clones were picked, the remaining cells that had been subjected to the DN-hTERT transfection, were harvested and passaged every 7 days to investigate the behaviour of the polyclonal mixed population (MP) during crisis and escape. These cells did not appear to go through crisis, as seen with the steady growth curve (figure 4.17 A) and the morphology of the cells which remained small and healthy overall. Furthermore, investigating the cell viability at each passage point showed no changes suggesting low levels of cell death between each sampling point (figure 4.17 A). This is most likely due to the fact that not all cells will be expressing the DN-hTERT which in turn dilutes cells that are entering crisis.

Due to the polyclonal nature of this culture, the STELA profiles showed no distinct homogeneous clonal populations. There was a mean erosion rate of 14 bps/PD and 11 bps/PD at the 17p and XpYp telomeres respectively between PDs 6 and 27, although not statistically different (P value of 0.34 and 0.56 for 17p and XpYp respectively following Kruskal-Wallis test). This was followed by a maintenance of the telomere length (figure 4.17 B and C) consistent with telomerase upregulation which was confirmed using the TRAP assay. The MP was also negative for C-circles (figure 4.17 D).

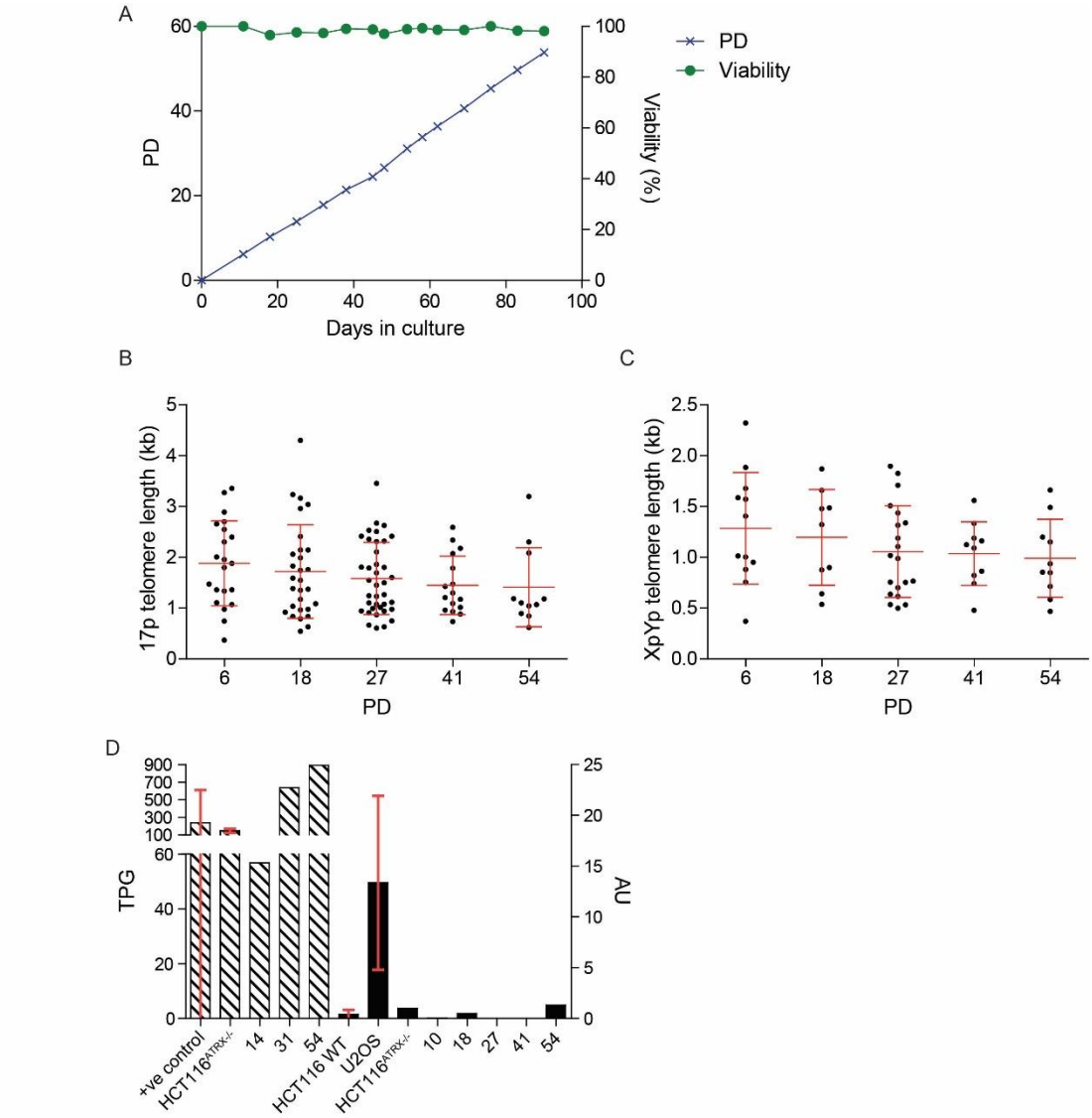


Figure 4.17: Characteristics of the HCT116$^{ATRX-/-\ DN\text{-}hTERT}$ mixed population.
A) Graph showing the dynamics of the cells in culture with the growth curve in blue, PDs depicted on left Y-axis, and the viability in green, percentage of viability depicted on right Y-axis, over the 90 days of cell culturing. Scatter plot of the telomere lengths measured by STELA at the B) 17p chromosome end and C) XpYp chromosome end over 5 sampling points. D) Graph showing the quantification of the total product generated (TPG) indicative of telomerase activity on the left axis and the C-circle quantification normalised to the HCT116$^{ATRX-/-}$ parental on the right axis expressed in arbitrary unit (AU). SD as error bars where possible.

4.4.8 C-Gas STING pathway

As previously mentioned, the STING protein has been shown to be downregulated in ALT cell lines (Chen et al. 2017). All HCT116$^{ATRX-/-\ DN\text{-}hTERT}$ ALT clones were therefore tested for the expression of the STING protein by Western blot. According to the above study, clones that upregulated ALT and died were expected to have high levels of STING contributing to high cell death, whilst clones that escaped crisis were expected to have low or absent STING expression, thus allowing cells to proliferate despite the accumulation of C-circles.

The results are presented in figure 4.18 and four categories are represented: escapee and ALT negative; no survival and ALT negative; no survival and ALT positive; escapee and ALT positive. The HCT116 WT and HCT116$^{ATRX-/-}$ parental cell lines showed similar STING expression and were used as controls for normal STING activity. The escapees that were negative for ALT had comparable STING activity to the controls as expected. The clones that were ALT negative and died had a downregulated STING expression suggesting that cell death was not mediated by the cGAS-STING pathway in these clones but instead due to crisis.

Variability was observed between no survival ALT positive clones. Indeed, clones that showed a comparable (11, 37, 51, 105, 108 and 128) or increased (67, 84 and 132) intensity of STING expression to the parental cells, consistent with the view that the accumulation of C-circles in the cytoplasm resulted in the initiation of the cGAS-STING pathway. In contrast, clones 20, 52, 74, 81 and 83 showed a downregulation of STING suggesting a dysfunctional DNA sensing pathway. This was however was not associated with continued cellular division implying that the loss of STING might be involved in the initiation of ALT and facilitate immortalisation and that other underlying events are required to achieve replicative immortality using ALT.

In the escapees that were ALT positive, clones 2 and 3 showed a slight downregulation of STING immediately before crisis (PD 20), perhaps suggesting that this step may be required early in the process of ALT thus enabling cells to divide despite the generation of C-circles observed by PD 30. The STING protein expression appeared to reach comparable levels to the controls upon upregulation of telomerase, suggesting the pathway was impaired temporarily. Furthering this hypothesis, clones 4, 48 and 131 showed comparable levels to the controls and these samples were taken after telomerase being upregulated as previous samples were unavailable.

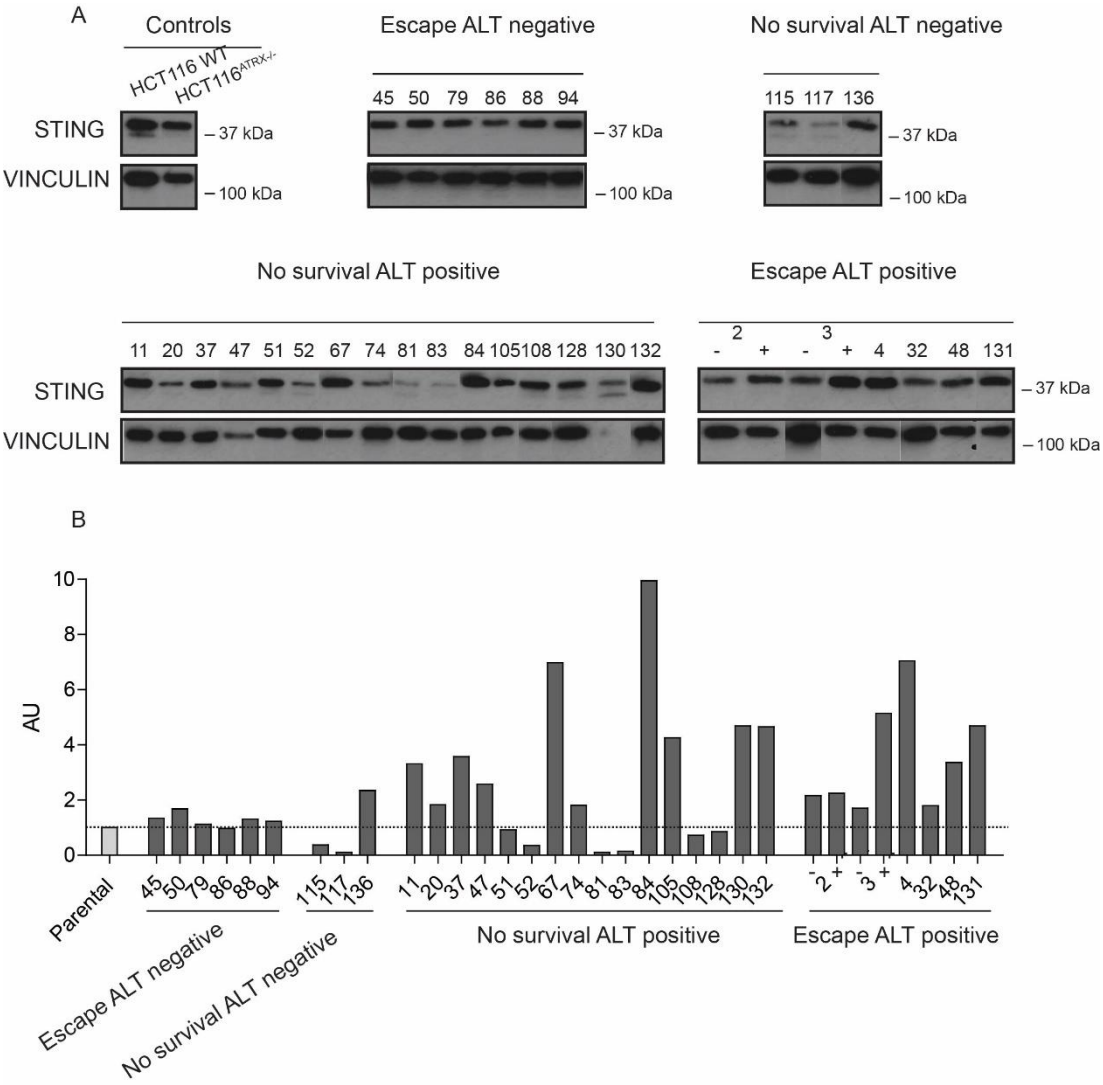


Figure 4.18: STING protein expression in HCT116$^{ATRX-/-\ DN\text{-}hTERT}$ clones.
A) Five panels representing the categories of clones tested: HCT116 WT and HCT116$^{ATRX-/-}$ parental cell lines as controls; escapee and ALT negative; no survival and ALT negative; no survival and ALT positive; escape and ALT positive. Vinculin used as loading control. B) Western blot quantification of STING expression (STING : Vinculin ratio) relative to the parental expressed in arbitrary unit (AU). The '-' representing before crisis and '+' after crisis. The numbers represent the clone number.

4.5 Discussion

4.5.1 The loss of ATRX combined with telomere stress affects cells' ability to escape crisis

The knockout of the *ATRX* gene in telomerase positive epithelial cancer cells and normal cells was not sufficient to initiate ALT activity, which provides evidence that the upregulation of telomerase is independent from ALT activation and represses this mechanism (Perrem et al. 1999; Napier et al. 2015). In contrast, combining the loss of ATRX with a telomere-driven crisis was enough to initiate ALT in escaping clones, as well as in clones that did not survive, resulting in ALT positivity in 19% of clones. In addition, this data showed that the loss of ATRX hindered the cells' ability to escape crisis as only 22% of clones survived as opposed to 100% of HCT116 WT clones which were taken through the same process and survived (Jones et al. 2014). Telomeres are difficult sites to replicate due to their repetitive sequence and the lack of ATRX is associated with increased replication fork stalling and collapse due to the formation of secondary structures, rendering replication in ALT positive cells more challenging (Wang et al. 2019). Cells lacking ATRX are therefore more prone to DNA damage and apoptosis which could explain the higher incidence of cell death in the HCT116$^{ATRX-/-}$ cells.

This further confirms the important role that ATRX plays at telomeres and the increased sensitivity to damage associated with replication fork stalling and collapse in these regions, explaining the use of HDR in ALT cells, a repair mechanism commonly used for resolving replication stress. Replication stress is thought to be part of the ALT phenotype and tightly controlled as replication fork resolution using BIR generates C-circles which contribute to the ALT phenotype and appear early in the initiation of the mechanism as they are thought to provide a template for telomere elongation (Zhang et al. 2019a; Zhang et al. 2019b). In addition, an enrichment of SMARCAL1 is observed at ALT telomeres to resolve replication barriers such as R-loops, limiting the amount of replication-associated damage to enable appropriate repair (Cox et al. 2016). In accordance with this, a stabilisation of secondary structures results in cell death due to an increase in unresolved barriers to replication, therefore, ALT cells have a tight control over replication stress (Rizzo et al. 2009). In addition, the presence of HDR related proteins present at telomeres and in APBs in ALT cells further suggests the importance of HDR-mediated repair in these cells (Yeager et al. 1999; Henson et al. 2002; Ait Saada et al. 2018).

4.5.2 ALT affects multiple chromosome ends and is chromosome-specific

The mechanisms of telomeric elongation in ALT cells remain unclear, however, evidence that HDR-based pathways are used to elongate telomeres for survival in ALT positive cells has been reported in the literature (Dunham et al. 2000; Teng et al. 2000; Nabetani and Ishikawa 2011). The data presented in this chapter provide evidence that ALT is chromosome specific due to a consistent insertion length for each chromosome end across independent ALT positive clones. The reason for this is unknown, nonetheless, one hypothesis could be related to the location of the individual chromosomes in tightly packed chromatin in the nucleus. Indeed, HDR occurs after S phase, when the chromatin is condensed, and an assumption could be that chromosomes that are located close to each other could be used as templates, suggesting inter-chromosomal recombination (Varley et al. 2002). Consistent with this hypothesis, a study showed that telomeric sister chromatid exchanges result in sequence heterogeneity but not in length heterogeneity, whereas inter-chromosomal exchanges affect telomere length suggesting that for successful telomere elongation multiple pairs of chromosomes need to be used as template (Liu et al. 2018).

The data also showed that ALT affects all short telomeres, in this case five separate chromosome ends, and that following elongation, telomeres erode as a function of cell division. As clones were unable to maintain ALT on the long-term in this model, it is not possible to address whether telomeres continually erode and undergo sudden elongation upon reaching a specific length as suggested in the literature (Murnane et al. 1994; Teng et al. 2000). However, telomere erosion results in background levels of genomic instability and dysfunctional telomeres throughout proliferation and a wide range of telomere lengths are observed in ALT cell lines, as opposed to telomerase upregulation, which stabilises the telomeres and minimises instability following crisis (Lovejoy et al. 2012; Jones et al. 2014; Li et al. 2019). Dysfunctional telomeres can once more be associated with the loss of ATRX and the impact this has on telomeric chromatin. Indeed, as previously mentioned, the lack of H3.3 incorporation through ATRX action results in a less condensed chromatin which exposes telomeres to the DDR and provides an explanation for the background level of instability, providing an explanation for the inability of a subset ALT positive clones (21) to survive.

4.5.3 ALT is a reproducible event

Cells that previously escaped crisis using ALT were taken through a second crisis and successfully escaped using the ALT mechanism once more. This shows the reproducibility of this phenomenon and suggests ALT initiation occurs shortly after the addition of the DN-hTERT. The downregulation of telomerase could have an immediate effect on cells that lack ATRX and thus telomeric elongation events may occur prior to the first sample. Furthermore, this hypothesis would also account for the clones that died despite being positive for ALT and for which no elongation was observed. The successful ALT-mediated escape of clones 2a, 3a and 4a could also be related to the location of the integration sites of the virus as the infection was not repeated in these clones and these were therefore somewhat expected to immortalise based on the observations made the first time the clones went through crisis.

Another aspect of reproducibility is proving that by repeating the infection three times, more ALT escapees were obtained. Evidence of ALT in non-surviving clones also showed the large-scale initiation of this mechanism despite it not being maintained. The initiation of ALT is seen at a high incidence, 19% overall and 4.7% of clones that survived, despite the rate of ALT positivity in epithelial cancer cells being low (0.9%) (Heaphy et al. 2011b). This could imply that ALT occurs through silencing or downregulation of the telomerase pathway although the reason for this is unclear (Atkinson et al. 2005). Indeed, in cells of mesenchymal origin, which readily upregulate ALT for survival, low levels of telomerase are observed during development and a repression of the catalytic subunit hTERT, preventing telomerase action, is common in these cells therefore suggesting a preference for activation of ALT for survival (Parsch et al. 2004; Lafferty-Whyte et al. 2009).

4.5.4 ALT cannot be maintained in epithelial cells

The HCT116$^{ATRX-/-}$ cells are epithelial and are positive for telomerase. Epithelial cells preferentially upregulate telomerase to maintain their telomere lengths explaining the much higher incidence of telomerase upregulation versus ALT overall (Heaphy et al. 2011b). Some studies suggest it is due to the developmental origin of cells in which mesenchymal stem cells require very little to no telomerase during development, as opposed to all other stem cells, thus accounting for the high incidence of ALT in cells of mesenchymal origin (Henson et al. 2002; Parsch et al. 2004).

The DN-hTERT downregulates telomerase by transporting the catalytic subunit hTERT into the cytoplasm and degrading it (Nguyen et al. 2009). A small percentage of clones (4.7%) were able to transiently upregulate ALT to enable further cellular division but ultimately, for long-term replicative survival, these cells require a re-activation of telomerase. These cells do not appear to have the ability to sustain ALT, but a knock out of the *hTERT* gene could address this as cells would be forced to maintain ALT for survival (La et al. 2016). This is however not a physiologically relevant method as telomerase is simply silenced in normal tissue following development and would be less feasible for therapy (Cong et al. 2002). Indeed, a decrease in hTERT expression, notably through increased expression of repressors of telomerase such as E2F1, in ALT cells has been shown to be implicated in maintenance of the ALT mechanism, as ALT cells do not activate the enzyme for survival or switch to this TMM (Lafferty-Whyte et al. 2009). Furthermore, the HCT116 cells have an intact *TP53* gene and ALT cells often appear to have a mutation in that gene, thus potentially suggesting it may facilitate upregulation and the maintenance of the ALT mechanism (Mirabello et al. 2015; Kratz et al. 2017; Oppel et al. 2019).

In addition, amongst the clones that were positive for C-circles and died, those with short telomeres appear to show less ALT activity than ones with long telomeres suggesting that elongation is required for a maintenance of ALT, even on the short-term. It also shows that C-circles begin to be generated early in the process of ALT initiation. This shows that these cells have the ability to initiate the ALT mechanism but, as previously mentioned, they require telomerase upregulation to survive on the long-term. This also suggests the implication of other unknown underlying events involved in the upregulation and maintenance of ALT as well as the elongation of short telomeres in this model. Furthermore, the STELA blots presented are only indicative of the XpYp telomeres and do not reflect the telomere dynamics at all other chromosome ends and a broader characterisation of telomere dynamics may be informative, although the sensitivity of STELA enabled the study of individual telomeres which showed a chromosome-specific elongation.

4.5.5 Co-existence of ALT and telomerase

The data collected show that when C-circles are at their maximal intensity when telomerase is absent. A study has shown that re-expressing ATRX in ALT positive cells inhibits ALT and a normal phenotype is restored (Clynes et al. 2015). Another study showed that low levels of telomerase is enough to inhibit ALT and ensure survival exclusively through the action of

telomerase in yeast (Millet et al. 2015). This data however shows a co-existence of ALT and telomerase. Maximal intensity of C-circles and telomerase do not coincide, but a strong C-circle positivity is observed by the last sampling points in some clones, even after a prolonged period of telomerase activation. Over the long-term, these C-circles may disappear as a function of cell division however, there is no evidence of dilution of the C-circles in this data despite prolonged cell culture and a rapid cell growth following crisis. This could imply that both telomerase and ALT are maintaining telomere length and competing for short telomeres (Perrem et al. 2001; Pompili et al. 2017). Another hypothesis could be the presence of a polyclonal population with some cells expressing telomerase whilst some proliferate by telomere elongation using the ALT mechanism. One population could potentially outgrow the other over a long period of time, as telomerase is thought to be a repressor of ALT activity, but this therefore suggests the co-existence of two populations utilising different TMMs rather than the co-existence of both mechanism within one cell (Perrem et al. 1999). This does however underline the complexity and heterogeneity in ALT activity and the stratification of patients according to TMM may not be as distinctive as anticipated.

4.5.6 Important consideration for therapy

Telomerase targeted therapies have gained a lot of interest recently as these could potentially target 85-90% of cancers (Harley 2008). Many avenues are being investigated such as downregulation of the catalytic subunit by binding to the active site preventing it from binding to the hTERC template (Imetelstat) (Roth et al. 2010). This approach would work in a similar way to the DN-hTERT used in these experiments in that the gene coding for hTERT is intact and the protein is continually being transcribed and translated. This data however shows evidence that a downregulation of telomerase can initiate ALT in 19% of cases and enabled prolonged cellular division in 4.7% of cases. This is an important consideration and a better understanding of the ALT mechanism is necessary to prevent the occurrence of a switch in TMM. In addition, ALT positive cancers are more aggressive and associated with a lot of instability throughout malignant proliferation which could further threaten these therapies (Lovejoy et al. 2012). Prognosis is also worsened in patients that have ALT positive cancer cells rather than telomerase (Matsuo et al. 2010; Lee et al. 2012). Furthermore, by understanding ALT, a novel therapy for ALT positive cancers could also be developed.

Using an approach that targets telomeres and does not directly interfere with telomerase activity could be a way of treating cells that have upregulated either TMM. One approach is

to elicit the DDR through the accumulation of G-rich oligonucleotides which increase the rate of γ-H2AX and cell death (Tsolou et al. 2008). This combined with an ATRX mutation could induce a high rate of apoptosis especially as ALT cells are more prone to damage. This could be a better approach in the context of telomerase cells being able to transition to ALT as they would be more susceptible to death. This does however highlight the need to understand the mechanism of ALT better, and more precisely the exact mechanism to elongate short telomeres as well as the template used, to develop a therapy to target tumours that utilise this mechanism for survival and hinder the pathway.

4.5.7 The involvement of other pathways in the maintenance of ALT

ATRX inhibits ALT in normal and telomerase positive cells and its loss appears to be involved in the initiation of ALT (Perrem et al. 1999; Brosnan-Cashman et al. 2018). However, cells that were unable to survive despite being positive for C-circles provide evidence that other underlying mechanisms are required for the maintenance of ALT. The involvement of another pathway required for the maintenance of this mechanism rather than the initiation is yet to be found as clones that died were able to initiate the generation of C-circles despite STING being active. This also shows that C-circles are involved very early in the ALT mechanism suggesting their potential role as DNA template for telomere elongation through RCA which would account for the sudden increase in telomere length associated with elongation using ALT (Teng et al. 2000).

In addition, the cGAS-STING pathway did not show a clear involvement in the ALT mechanism in the HCT116 model and did not appear to be required for ALT upregulation, despite a complete loss of the STING protein in ALT positive cancer cell lines alongside ALT-positive transformed cell lines in a recently published study (Chen et al. 2017). The reason for this discrepancy is unclear although the fact that these cells only transiently upregulate the ALT mechanism followed by telomerase upregulation could be a reason for these observations. This also suggests that the prolonged impairment of this pathway is required for maintenance of the ALT mechanism as these cells were unable to maintain it, as seen with levels of STING protein expression levels comparable to the parental cell line by the last sampling point.

Downregulation of ATRX as well as the DAXX and H3.3 proteins has been associated with ALT and a mutation in either of these three proteins is enough to prevent the incorporation of the H3.3 histones within telomeres and therefore cause alterations to the chromatin

structure (Lewis et al. 2010; Episkopou et al. 2014). ATRX is more commonly mutated and its loss is consistently seen in the majority of ALT positive cells, presumably due to its wider range of activity and the role it plays in a DAXX-independent way, such as in sister chromatid cohesion (Ramamoorthy and Smith 2015). Some tumours do however retain WT ATRX activity which suggests the involvement of another unknown pathway in the mechanism of ALT.

4.5.8 Conclusions

In conclusion, the loss alone of ATRX does not affect telomerase positive cells. In contrast, the combination of a telomere-driven crisis and the loss of ATRX reduces the rate at which epithelial cells readily escape crisis when compared to WT cells. In addition, the initiation of ALT was observed in 21% of escapees (4.7% overall) transiently and 18% in clones that died (14% overall). The ALT mechanism was not maintained in any clone as seen with the upregulation of telomerase for long-term survival in escapees and the death of all cells in clones that were unable to re-activate telomerase. The mechanisms underlying ALT initiation and maintenance remain largely unknown and the ability of cells to re-activate telomerase in comparison to others is not clear. This does however show that epithelial cells with inhibited telomerase are able to switch to ALT for survival, but implies the involvement of other underlying mechanisms involved in the maintenance of the ALT mechanism.

Chapter 5

Understanding the mechanisms that underline ALT and their impact on the cancer genome

5.1 Abstract

The escape from a telomere-driven crisis following the re-establishment of telomerase activity is associated with genomic rearrangements, however, the extent to which this occurs in cells that upregulate ALT is not clear. To investigate the genomic impact of the upregulation of ALT during crisis, whole genome sequencing (WGS) and single-molecular real-time (SMRT) long-range sequencing of telomeres were undertaken.

Telomeres from ALT positive cells HCA2$^{HPV\ E6E7\ ATRX-/-}$ clone 21 (chapter 3), HCT116$^{ATRX-/-\ DN-hTERT}$ clone 2 (chapter 4) and U2OS were sequenced using the PacBio SMRT sequencing platform. These were supplemented with the HCA2$^{HPV\ E6E7}$ parental and HCT116$^{ATRX-/-}$ parental used as matched normal ALT negative controls. The data was filtered and analysed for ALT-specific features. Upon analysis, it was made obvious that an enrichment in TTCGGG variant repeat was observed in the HCT116 model, whilst an overall increase in different variant repeats was observed in the HCA2 and U2OS models. These data indicated that this phenomenon is cell specific and no increase in a particular telomere repeat variant was associated with ALT, but rather overall increase in non-canonical repeats was observed. In addition, the WGS data obtained from HCT116$^{ATRX-/-\ DN-hTERT}$ clones pre- and post-crisis following a telomerase or ALT upregulation, showed the surprising lack of structural variants (SV) in ALT clones, compared to telomerase positive clones. This is suggestive of a more stable genome in ALT clones, despite evidence of chromothripsis in one ALT positive clone.

In conclusion, in the early stages of ALT, the genome appears to be more stable compared to telomerase positive cells, whilst the telomere sequence differs greatly between the parental and their respective ALT clone thus indicating an increase in variant repeat content as a result of ALT elongation.

5.2 Introduction

Telomere crisis is associated with a high incidence of chromosome instability and subsequently cell death. During this phase, critically short telomeres can no longer protect the natural ends of chromosomes from the repair machinery resulting in aberrant end-to-end fusions (Counter et al. 1992; Capper et al. 2007). This leads to mis-segregation and tearing of chromosomes during cellular division thus resulting in aneuploidy and complex karyotypes associated with some cancers (Bayani et al. 2007). By this means, DNA harbours structural variants (SVs) and copy number variants (CNVs) that will sustain upon escape from crisis and these include breakpoints, translocations, duplications, inversions and deletions. Next-generation sequencing (NGS) of DNA allows for the distinction of these complex rearrangements that arise from cells transiting through crisis (Guan and Sung 2016; Cleal et al. 2019).

In some cases, (2-3%), a phenomenon termed chromothripsis can also arise during crisis (Stephens et al. 2011). Chromothripsis is a catastrophic event during which chromosomes suffer intense localised rearrangements, independently of NHEJ, resulting in highly rearranged genomes that can involve multiple chromosomes (Cleal et al. 2019). ALT positive cancer cells have been reported to be genetically unstable and this phenomenon occurs more frequently in some commonly ALT positive cancers such as osteosarcoma, glioblastoma and up to 100% of liposarcoma tumours tested showed signs of chromothripsis (Stephens et al. 2011; Cortes-Ciriano et al. 2018). More importantly, it often involves numerous chromosomes resulting in more heterogeneity which is confirmed by the complexity of the karyotypes observed in cancer and especially sarcomas (Guillou and Aurias 2010). The ALT status was not reported in these studies, however, as NHEJ is silenced in the context of ALT upregulation, the ALT mechanism could potentially increase the rate of chromothripsis, although this has not been assessed as of yet (Koschmann et al. 2016).

In addition, studies have shown that ALT telomeres differ from telomeres elongated by telomerase action upon analysis of the telomere repeat array, thus providing a new hallmark for diagnosis of ALT positive cancer cells (Varley et al. 2002). An overall increase in most variant repeats was observed upon comparison between ALT positive and negative cells resulting in a clear difference between telomerase and ALT elongation (Lee et al. 2014). A cell specific increase in certain variants was also apparent in PanNET and melanoma cells (Lee et al. 2018). This alteration in variant repeats showed evidence of reduced binding of the Shelterin complex, which exclusively binds to TTAGGG canonical repeats, allowing for novel

protein interactions with the ends of chromosomes (Court et al. 2005; Conomos et al. 2012). This in turn may play a role in the ALT phenotype through the increase of HDR-mediated elongation of telomeres (Baumann and Cech 2001; Court et al. 2005). It has been suggested that this variability in variant repeats and the presence of those throughout the telomere length arises from recombination events between short and long telomeres. Indeed, it is a well-established fact that human telomeres have acquired variant repeat patterns at the start of the telomere, up to 1 kb, thus providing heterogeneity in repeat content (Allshire et al. 1989; Baird et al. 1995; Baird et al. 2000).

Overall, the DNA sequence of ALT positive cancer cells appears to differ from normal and telomerase positive cells, and sequencing of DNA from ALT cells could potentially reveal ALT-specific characteristics and provide details on the mechanism of ALT elongation through the study of individual telomeres.

5.3 Aims

The overall hypothesis for this chapter is that the initiation of ALT induces alterations in the telomere sequence and may have an impact on the overall cancer genome. To address this, telomeres from three different cell types, comprising of ALT positive and controls, were sequenced to examine the telomere sequence following ALT upregulation and survival. In addition, WGS was carried out to further address the possible effects of the loss of ATRX combined with telomere stress on the overall genome. The aims for this chapter are as follows:

- Sequence telomeres using the PacBio platform to obtain full long telomere reads
- Characterise the elongation event in ALT telomeres
- Analyse the telomere sequences to establish if ALT occurs as a single event of if it associated with multiple events
- Assess if there is a common event across three different cell types or if ALT elongation is cell specific
- Establish if ALT affects all alleles and specifically short telomeres
- Determine the length at which telomeres become elongated in ALT cells
- Establish any effects that loss of ATRX has on the genome as cells transit through crisis and upregulate ALT or telomerase
- Assess if there are any differences in impact on the genome between clones that upregulated ALT and died; upregulated ALT and survived; and upregulated telomerase and survived

5.4 Results

5.4.1 Sequencing of ALT positive telomeres

5.4.1.1 Choice of sequencing platform

The Sequel single-molecular real-time (SMRT) sequencing by Pacific Biosciences (PacBio) was used to sequence specific telomeres. This sequencing platform allows for the generation of long reads (over 20 kb with an average of 10-14 kb), with therefore, the potential to obtain whole telomere sequences generated using STELA, between the sub-telomere primer and the telorette primer at the chromosomal terminus in one read, as opposed to attempting to assemble repetitive contigs generated from Sanger sequencing. SMRT sequencing generates highly accurate reads (over 99%); presents less issues with GC rich regions such as telomeres; and allows for single-molecule resolution which would enable the study of single events (Eid et al. 2009).

Details of the library preparation are presented in figure 5.1 and consist of a DNA repair step to generate blunt ends to enable the ligation of SMRTbell adapters to the ends of amplicons which in turn allows the annealing of the primer and polymerase. The amplicons are then loaded onto a SMRT cell 1M V3 LR containing 1,000,000 zero-mode waveguides (ZMW) ensuring that a maximum of ZMWs only contain one molecule. This molecule is then anchored to the bottom of the well initiating DNA polymerase synthesis of a new DNA strand by incorporation fluorophore tagged dNTPs, which do not interfere with DNA polymerase action, and allows for the detection of a fluorescent signal specific to each base pair. This generates circular consensus sequencing reads (CCS) which consist of aligned subreads from a single ZMW which form a highly accurate single read (Ardui et al. 2018).

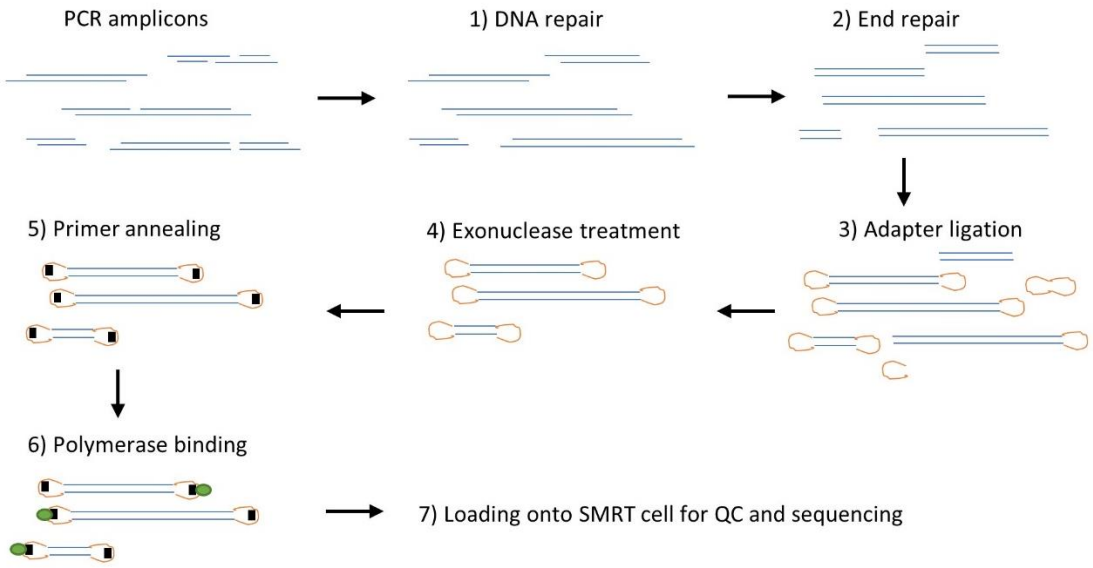


Figure 5.1: Schematic representation of PacBio library preparation of PCR amplicons prior to sequencing.
Double blue lines represent the PCR amplicons; orange semi-circles represent the SMRTbell adapters; black rectangles represent the primer; and green circles represent the DNA polymerase.

5.4.1.2 Selection of clones to be sequenced

A total of five samples were selected to be sequenced. Firstly, the HCT116$^{ATRX-/-}$ parental cell line (figure 5.2 A) was chosen to obtain pre-crisis reads with a background of telomerase elongation to be used as a control and compared to the second sample, HCT116$^{ATRX-/-\ DN-hTERT}$ clone 2 (figure 5.2 B), which underwent ALT-like elongation, to establish any ALT-like characteristics in these cells. This clone was picked as it had undergone telomere elongation at multiple chromosome ends and sufficient DNA was available for a large-scale amplification of single telomeres. To limit the amount of input genomic DNA, 250 pg of DNA per PCR reaction were used, therefore dismissing 9p and 5p chromosome ends, which require a minimum of 2.5 and 1.25 ng of DNA per PCR reaction for telomere amplification using STELA. Therefore, the 17p, XpYp and 7q telomeres were amplified in a multiplex PCR reaction to further limit the input genomic DNA quantity, the output sequencing reads were to be separated and mapped bioinformatically to the appropriate sub-telomere region upon successful sequencing by Dr Kez Cleal.

In addition, the HCA2$^{HPV\ E6E7}$ parental fibroblast cell line (figure 5.2 C) was sequenced to provide pre-crisis reads associated with normal fibroblasts to be used as a control and compared to the fourth sample, HCA2$^{HPV\ E6E7\ ATRX-/-}$ clone 21 (figure 5.2 D), which successfully escaped crisis by upregulating ALT for survival. Clone 21 was picked as it showed the widest range of telomere lengths post-crisis (0.6-10.4 kb and 0.6-10.3 kb at the 17p and XpYp

chromosome ends respectively) compared to other ALT clones generated from the same experiment. 375 pg of DNA were used per PCR reaction and the 17p and XpYp telomeres were amplified in the same PCR.

Finally, the U2OS osteosarcoma cell line (figure 5.2 E), which is ALT-positive, was chosen to further address events associated with ALT and to assess the occurrence of ALT-characteristics across multiple cell types or alternatively, evidence that ALT is cell specific. 625 pg of DNA were used per PCR reaction and the 17p and XpYp telomeres were amplified in the same PCR.

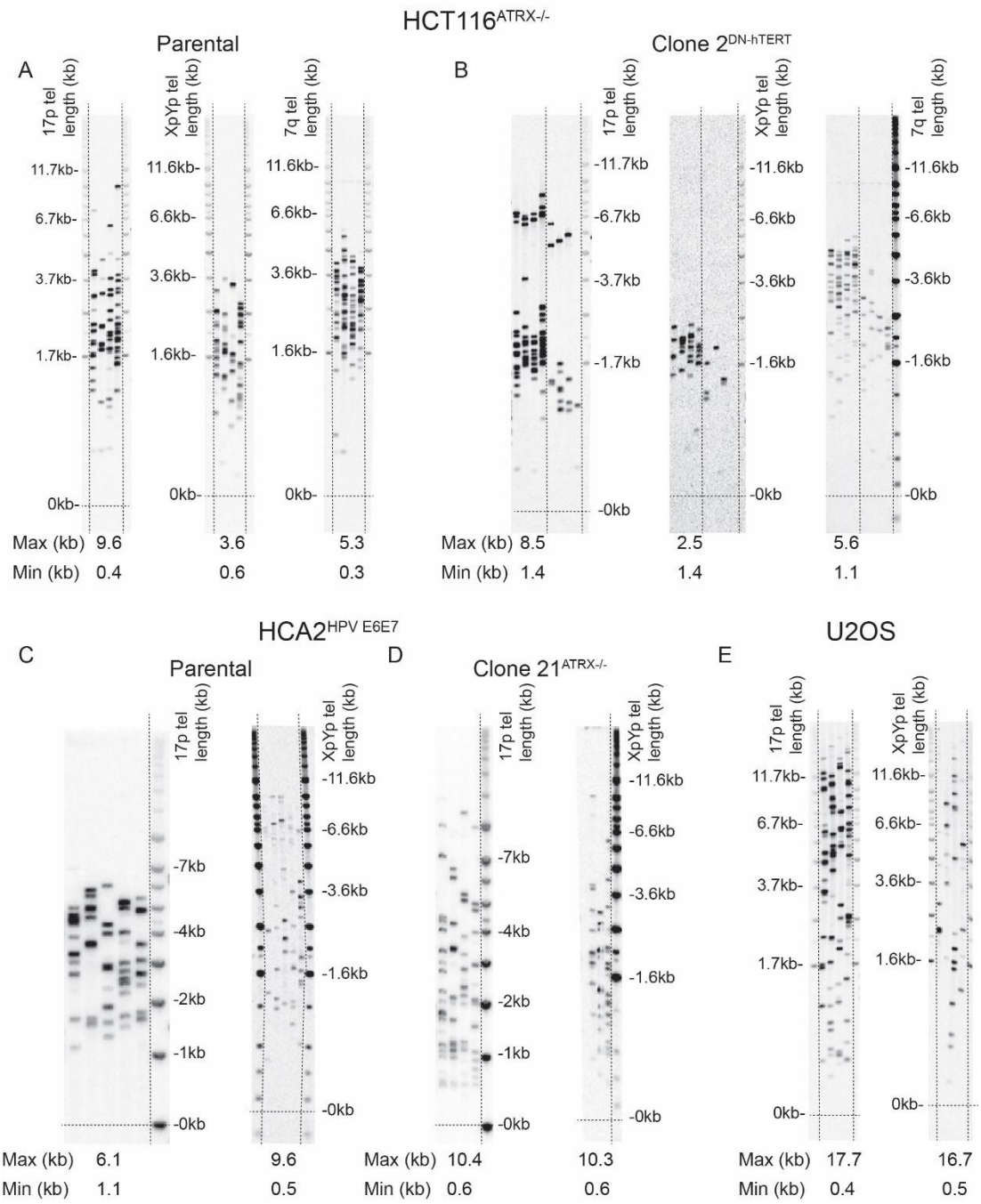


Figure 5.2: Samples picked for PacBio sequencing upon STELA analysis at multiple chromosome ends.

STELA profiles at the 17p, XpYp and 7q chromosome ends for A) HCT116$^{ATRX-/-}$ parental cell line; B) HCT116$^{ATRX-/-\ DN-hTERT}$ clone 2. STELA profiles at the 17p and XpYp chromosome ends for C) HCA2$^{HPV\ E6E7}$ parental cell line; D) HCA2$^{HPV\ E6E7\ ATRX-/-}$ clone 21; and E) U2OS. With the longest and shortest telomeres measured on each blot across the bottom labelled max and min expressed in kb.

5.4.1.3 Optimisation of PCR

An initial STELA PCR was done to establish the optimal number of PCR cycles to generate a sufficient number of molecules without saturating PCR reagents and inducing non-specific bands. It was concluded that 24 cycles would subsequently be used as a smaller quantity of telomeres were observed after 22 cycles and a non-specific band appeared at around 0.1 kb after 26 cycles (figure 5.3 A). Furthermore, due to the amplification of multiple telomeres by the inclusion of multiple primers in each PCR mix, for which different annealing temperatures are normally required (i.e. 59 °C for 17p primers and 65 °C for XpYp and 7q primers), a gradient PCR was performed for all samples to establish the optimal annealing temperature to yield the maximum number of molecules (figure 5.3 B-F). All samples showed the optimal results at 63 °C.

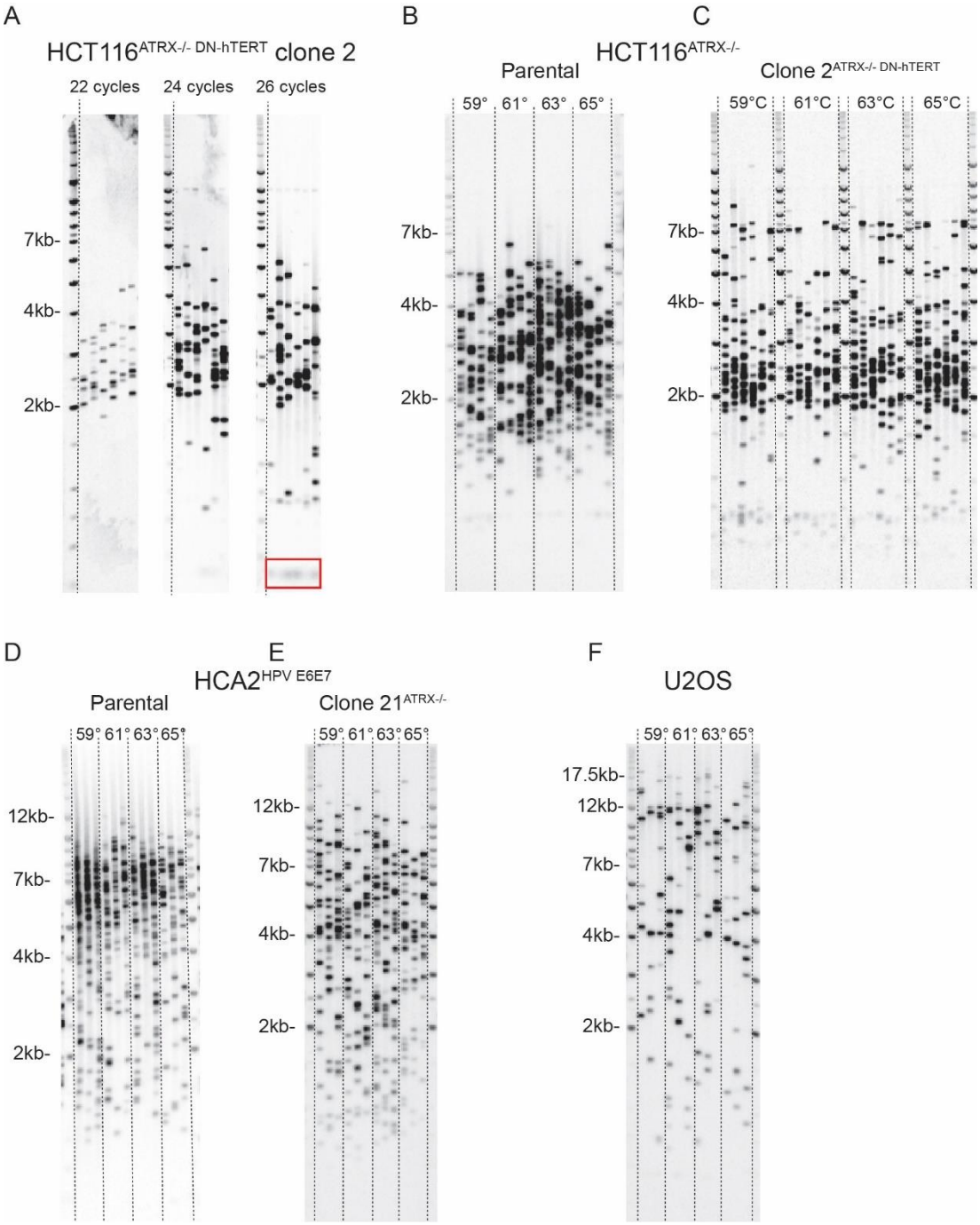


Figure 5.3: Optimisation of the PCR for optimal yield.
A) STELA profile resulting from 22, 24 and 26 PCR cycles for HCT116$^{ATRX-/-\ DN-hTERT}$ clone 2 with the red box highlighting unspecific bands; STELA profiles following a gradient PCR for B) HCT116$^{ATRX-/-}$ parental cell line; C) HCT116$^{ATRX-/-\ DN-hTERT}$ clone 2; D) HCA2$^{HPV\ E6E7}$ parental cell line; E) HCA2$^{HPV\ E6E7\ ATRX-/-}$ clone 21; and F) U2OS cell line.

5.4.1.4 Generation of sequencing reads

The PacBio sequencing protocol requires an initial DNA concentration, prior to library preparation, of 500 ng minimum. To achieve this, 1,600 PCR reactions were amplified for each sample. The quality control (QC) results prior to library preparation are presented in figure 5.4 and consisted of 4-5 STELA reactions to ensure appropriate amplification of telomeres (lanes 1); 0.03% of pooled PCR reactions total volume prior to SpeedVac volume concentration (lanes 2); 0.03% of pooled volume post Speedvac volume concentration diluted to 5µl to ensure no loss of DNA was incurred during volume concentration (lanes 3); 0.1% of total volume pre-AMPure purification (lanes 4); and finally 0.1% of total volume post-AMPure purification diluted to 5 µl to ensure no loss of DNA was incurred during purification step (lanes 5). Samples were then sent to be sequenced following appropriate library preparation by Mrs Joanne Morgan (NGS coordinator, Cardiff University).

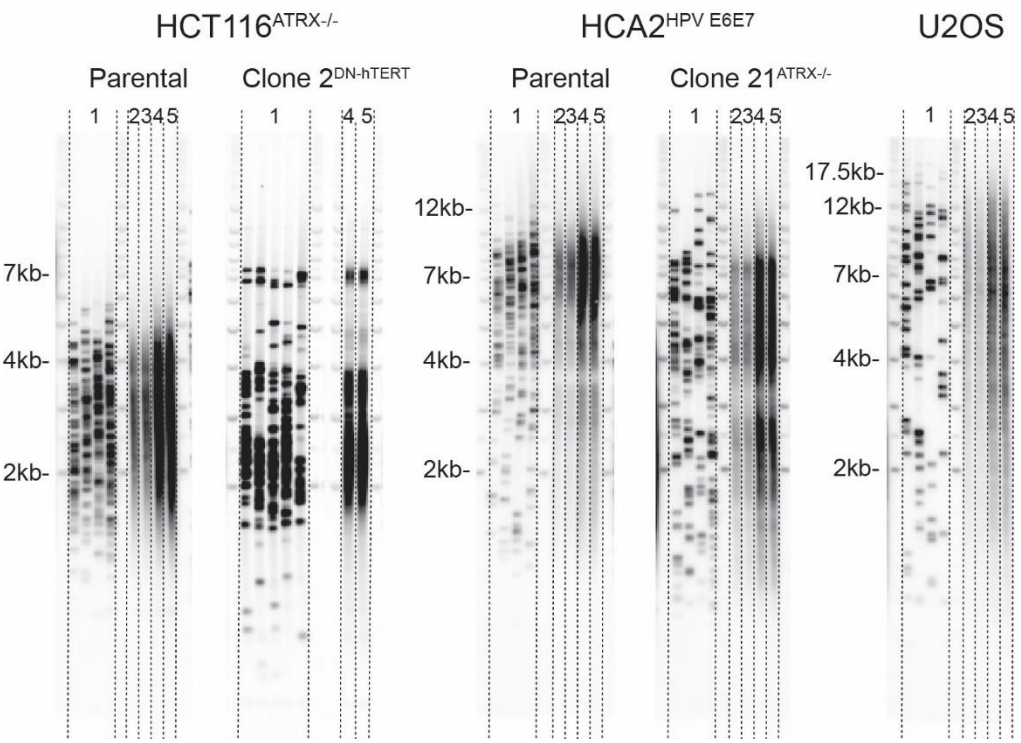


Figure 5.4: Quality control of samples sent for sequencing.
1: STELA reactions; 2: 0.03% volume of pooled PCR reactions (pre-Speedvac volume concentration treatment); 3: 0.03% post-Speedvac volume diluted to 5 µl; 4: 0.1% pre-purification volume; 5: 0.1% post-purification volume diluted to 5 µl.

5.4.1.5 Processing the data

The data was processed and filtered by Dr Kez Cleal. Firstly, the raw reads were filtered to retain only reads with a sub-telomere primer at one end and a telorette primer at the other end (Appendix 1). This was done by aligning the reads using Edlib. Then, all sequences were labelled using a hidden Markov Model (HMM) to highlight and dissociate telomere repeat arrays, sub-telomere sequences and insertions (figure 5.5 A). Sequences were broken down into 6 bp kmers and given scores upon comparison to the canonical telomere repeat TTAGGG allowing an edit distance of 2 bps. The scores were as follows: 0 for background sequence (sub-telomere and insertions); 1 for forward strand telomere repeats (CCCTAA); and 2 for reverse strand telomere repeats (TTAGGG) (Examples in Appendix 2). Therefore, variant repeats with a maximum of 2 bps substitutions compared to TTAGGG were classed as telomere sequence. To clean the data further, subsequent filtering steps were added to the pipeline which aimed at removing sequencing and PCR artefacts generated during the process. By this means, the following classes of reads were removed from the analysis: unexpected non sub-telomeric sequences amplified by low homology with primers; STELA products that appeared to have undergone primer swapping; products that did not have a detectable sub-telomere sequence; and concatemers of STELA products. All retained sequences were then compiled into an Excel spreadsheet for manual curation and analysis. The spreadsheet included the sub-telomere length, trimmed telomere length as well as the insertion lengths and sequences amongst other features. A schematic representation of a sequence is represented in figure 5.5 B.

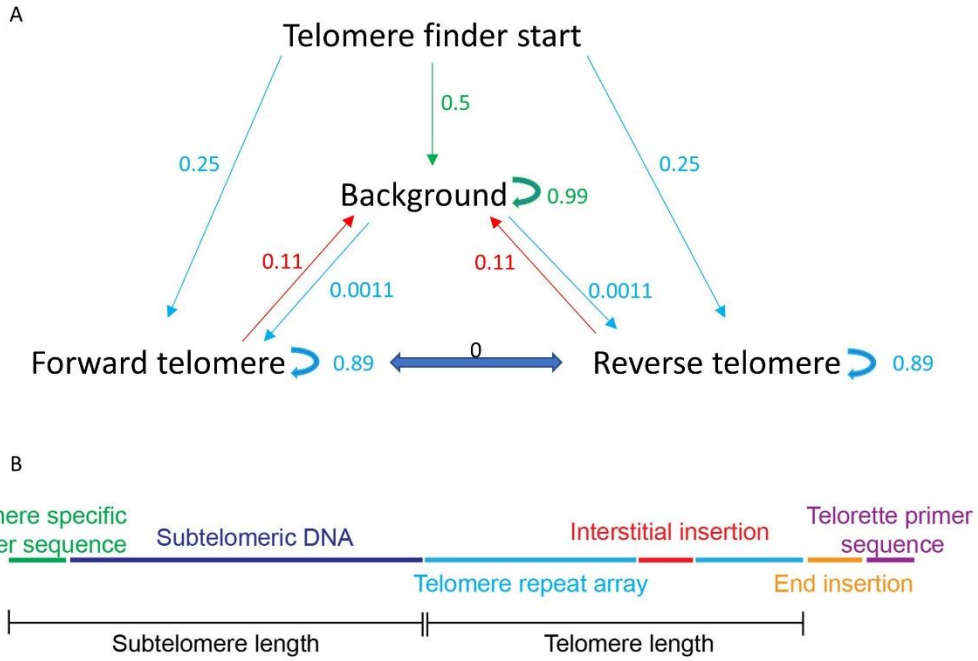


Figure 5.5: Schematic representation of a sequencing read used for analysis.
A) Probability of the Hidden Markov Model of obtaining background sequences (sub-telomere DNA; interstitial insertions; end insertions; and telorette primer); forward oriented telomere; or reverse oriented telomere. In green the probability of sub-telomere sequence; in blue the probability of telomere sequence; and in red the probability of an insertion sequence within the telomere. B) The sub-telomere length includes the telomere specific primer and the sub-telomeric DNA sequence. The telomere length includes the repeat array and any interstitial insertions. End insertions are not included in the telomere length similarly to the telorette primer sequence.

Sample	Primers	Reads primer + telorette	Average length (kb)	Range from sequencing (kb)	Range from STELA (kb)
HCT116$^{ATRX-/-}$ Parental	17pseq1rev	85,609	1.75	0.019 - 8.6	0.4 - 9.6
	XpYpC	35,204	1.53	0.025 - 5.4	0.6 - 3.6
	7qK1	22,157	2.34	0.026 - 6.0	0.32 - 5.3
HCT116$^{ATRX-/-}$ $^{DN-hTERT}$ Clone 2	17pseq1rev	33,003	1.26	0.017 - 6.2	1.4 - 8.5
	XpYpC	10,889	1.27	0.029 - 5.4	1.4 - 2.5
	7qK1	16,295	1.11	0.018 - 5.1	1.1 - 5.6
HCA2$^{HPV\ E6E7}$ Parental	17p6	2,704	2.18	0.061 - 6.7	1.1 - 6.1
	XpYpE2	48,578	1.27	0.019 - 10.3	0.48 - 9.6
HCA2$^{HPV\ E6E7}$ $^{ATRX-/-}$ Clone 21	17p6	47,642	1.02	0.019 - 8.6	0.57 - 10.4
	XpYpE2	85,748	1.15	0.018 - 10.5	0.57 - 10.3
U2OS	17pseq1rev	148,382	1.96	0.019 - 16.4	0.35 - 17.7
	XpYpE2	28,412	1.91	0.019 - 15.1	0.46 - 16.7

Table 5.1: Summary of telomeres amplified for each sample.
The total number of reads and the range of telomere lengths following sequencing and on the matching STELA blots.

5.4.2 Separation of alleles

The telomere length data obtained in HCT116 cells escaping crisis in the absence of ATRX, indicated that the shortest telomeric alleles were subjected to specific ALT lengthening events. To examine the allelic composition of telomeres before and after crisis, the telomere variant repeat patterns identified in the long read sequence data were used to discriminate between telomeric alleles that have previously been shown to display extensive allelic variation (Baird et al. 1995; Baird et al. 2000). To establish the occurrence of a variant content pattern, common variant repeats were colour coded as follows: TTAGGG; TCAGGG; TTCGGG; GTAGGG; TGAGGG; TTGGGG; TAAGGG; CTAGGG; TTTGGG; TGCGGG; AGAGGG. Each of those variants, if identified in the sequences, have been replaced by a square (□) to compress the reads. Letters that have not been highlighted are single nucleotides. This legend will be used throughout this chapter. For this analysis, only reads that had a matching primer and sub-telomere sequence were used.

5.4.2.1 Method

To assess if ALT is allele specific or if it affects both alleles at multiple chromosome ends, the alleles for each sample, apart from HCT116$^{ATRX-/-}$ samples at the XpYp chromosome end for which only one allele is present and amplified, were separated based on the variant repeat sequences within the first 100 bps of the trimmed telomere sequence. The length of 100 bps was decided for consistency across all samples and to increase the selection stringency between alleles therefore obtaining a more accurate result. Each sequence was subsequently given a cosine similarity score based on similarities in the nucleotide sequence, with 0 being 100% similar and 1 being 0% similar, allowing for the selection of a specific allele and this was performed by Dr Kez Cleal. The separation of the alleles and subsequent analysis was undertaken by myself, and the threshold of 0.10 was used to separate the alleles. This threshold was picked to maximise the total number of reads usable and to account for natural variation within these sequences. In example, in figure 5.6, the lowest (0.10) and highest (0.01) scoring reads for HCT116$^{ATRX-/-}$ parental cell line at the 17p chromosome end allele 1 are presented and these showed a clear distinction from the allele 2 sequence (score of 0.50 and 0.35 respectively). Upon comparison, differences within the first 100 bps were observed whilst the remaining sequence appeared largely similar. Examples of the highest scoring sequences for each sample at each chromosome end studied are included in appendix 3, all showing a different pattern of variant repeats within the first 100 bps. The U2OS XpYp

telomeres showed no obvious variant repeat pattern preventing the dissociation of two alleles if present.

Score of 0.10 for allele 1 and 0.50 for allele 2

AAGGGTTAGGGTTAGGGTTGGGGTTGTGGTGGTTTGGGGTTGGGGTTAGGCTTAGGGGCTTAGGGCTTAGG
GGCTAGGGCTATGGGCTAGGGGCTAGAGT

Score of 0.01 for allele 1 and 0.35 for allele 2

AAGGGTTAGGGTTAGGGTTGGGGTTGGGGTTGGGGTTGGGGTTAGGGCTTAGGGCTTAGGGCTTAGGGCTA
GGGCTAGGGCTAGGGCGAGAGTTAGGGTT

Figure 5.6: Examples of reads obtained upon allele separation based on the variant repeat pattern in the first 100 bps.
Both examples are taken from the HCT116$^{ATRX-/-}$ parental cell line at the 17p allele 1 chromosome end and only the first 100 bps are displayed. At the top, the lowest scoring read with 0.10 and at the bottom the highest scoring read with 0.01.

5.4.2.2 Allele lengths

Upon allele identification, the telomere lengths were plotted to establish any differences in length and address the hypothesis that the shorter allele at the 17p chromosome end is elongated in the HCT116 model and to highlight this phenomenon, if present, in the other models (figure 5.7). Due to the PacBio platform preferentially sequencing short reads, the proportion of long reads does not reflect the telomere length distributions observed with STELA. Furthermore, the HCT116$^{ATRX-/-}$ ALT clone was the first sample to be sequenced and was underloaded due to a software malfunction further reducing the number of successful long reads. Therefore, the mean telomere length obtained from the PacBio reads does not match the mean determined using STELA. For example, the average telomere length of the U2OS at the XpYp chromosome end is 3.1 kb (median 2.8 kb) following PacBio sequencing and 5.7 kb (median 4.7 kb) upon STELA. However, the range of telomere lengths is similar between the two with 0.019-15.1 kb for PacBio and 0.46-16.7 kb for STELA thus still providing a good breadth of reads and enough material to address the aims set out for this chapter.

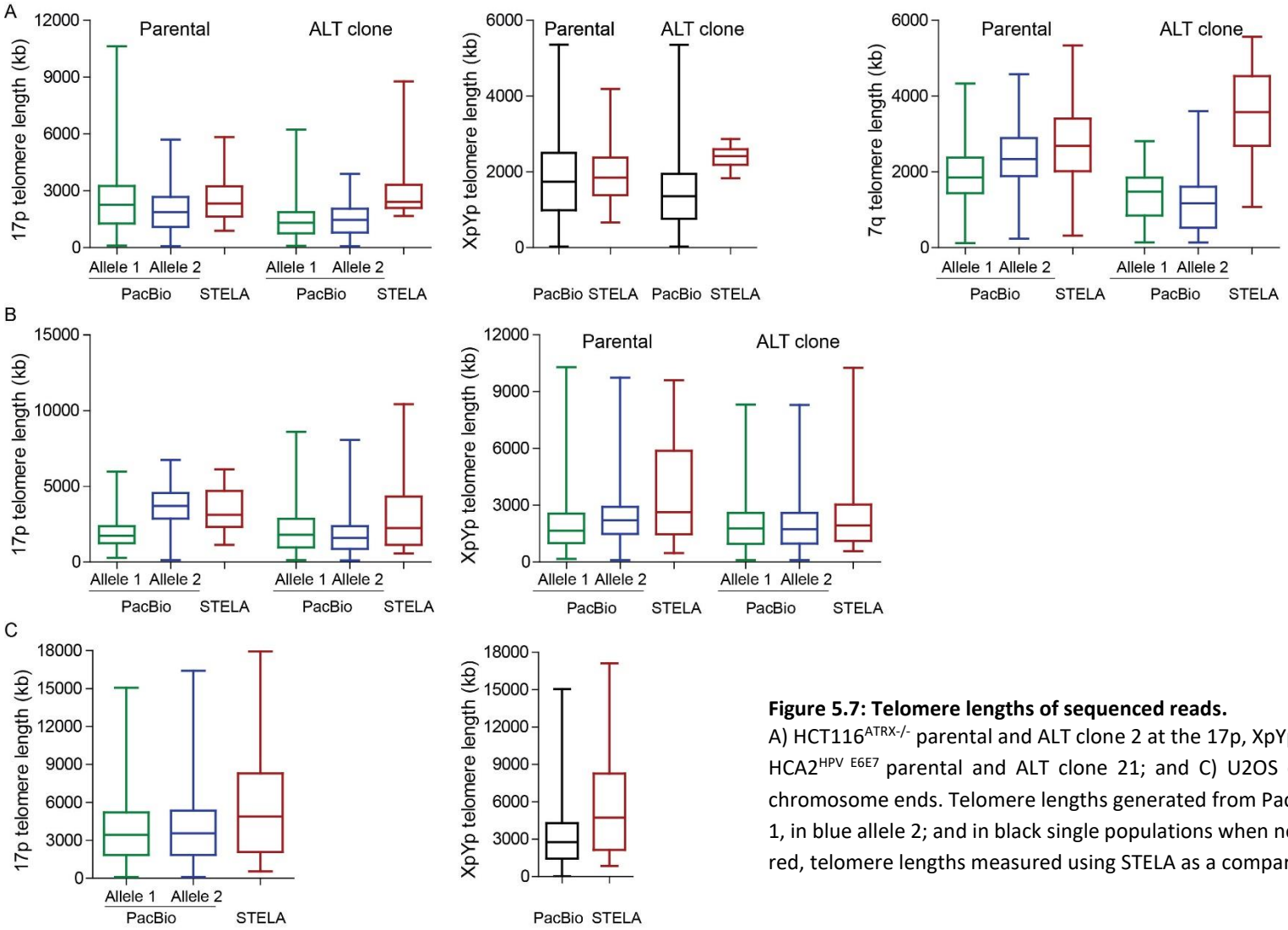


Figure 5.7: Telomere lengths of sequenced reads.
A) HCT116$^{ATRX-/-}$ parental and ALT clone 2 at the 17p, XpYp and 7q chromosome ends; B) HCA2$^{HPV\ E6E7}$ parental and ALT clone 21; and C) U2OS cell line at the 17p and XpYp chromosome ends. Telomere lengths generated from PacBio sequencing: in green allele 1, in blue allele 2; and in black single populations when no alleles were detected. In dark red, telomere lengths measured using STELA as a comparison.

5.4.3 Increase in variant repeats in ALT clones

The occurrence of each variant was calculated for each sequence at each chromosome end studied to highlight possible changes in telomere repeat content between the parental cell lines and their respective ALT clone. To further assess if the changes observed were specific to ALT, all ALT samples were compared to the U2OS cell line used as an ALT positive control.

5.4.3.1 HCT116 cell line

Allelic variant content differed between alleles in the parental cell line. A set of common variant repeats were counted and the overall proportion of each was plotted alongside the fold change in each variant when comparing the ALT clone to the parental to identify events associated with ALT elongation (figure 5.8). All alleles from different telomeres had a distinct variant repeat pattern, notably towards the start of the telomere thus enabling the separation of alleles, which was maintained between the parental and the ALT clone. However, the incidence of TTCGGG repeats consistently increased at all chromosome ends studied with up to a 550-fold increase at the XpYp chromosome end (figure 5.8 F). The TTAGGG repeat content was also consistently reduced in the ALT clone signifying an increase in non-canonical variant repeats when ALT is upregulated. Furthermore, the proportion of other variants (any variant that is not stated below) increased at all chromosome ends further confirming the overall increase in variant repeats that do not consist of TTAGGG.

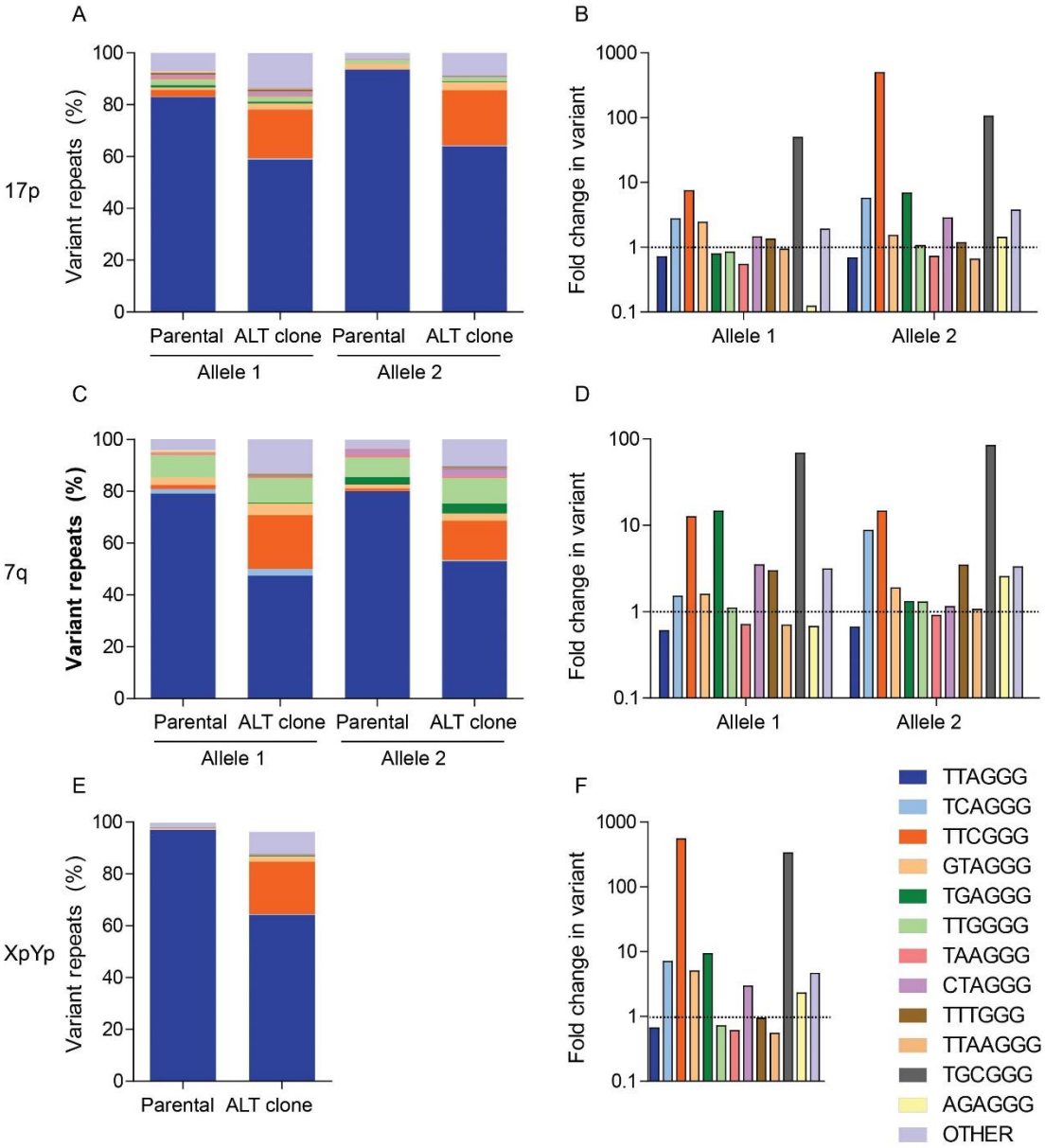


Figure 5.8: Telomere variant repeat proportions in HCT116$^{ATRX-/-}$ cells.
Total proportion of specific variant repeats combining all reads expressed in percentage for the parental and the ALT clone at the 17p (A); 7q (C); and XpYp (E) chromosome ends. Matching graph expressing the fold change in variant repeat proportion when comparing parental and ALT clone using a log scale for 17p (B); 7q (D); and XpYp (F).

To further confirm these findings, the common language effect size (CLES) indicator was applied (table 5.2). The CLES tests the probability of a random number from set B (proportion of a given variant in a random read from the ALT clone) being larger than a random number in set A (proportion of same variant in a random read from the parental) (McGraw and Wong 1992). A probability below 0.2 is thought to have no effect; between 0.2 and 0.49 a small effect; between 0.5 and 0.79 a medium effect; and above 0.8 a large effect. A value of 0.5 means that there is a 50% chance of a random number of set B being larger than a random

number in set A. The proportion of a given variant was normalised to the overall length of each read to establish a proportion of each variant relative to the length of the telomere.

On this basis, the TTAGGG repeat content was not significantly reduced in the ALT clone whilst the TTCGGG content presented a large effect size at all chromosome ends and confirmed observations made in figure 5.8. Another variant repeat that appeared increased in the ALT clone was TGCGGG (figure 5.8) however, a low, to no effect size was observed following the CLES test and can be attributed to the fact that this variant frequency was low in the parental and any increase would affect the fold change. In addition, when combining the proportion of all variant repeats, excluding TTAGGG (termed ALL in table 5.2), the results presented a large effect size at all chromosome ends further showing an increase in variant repeats in the ALT clone.

	17p		**7q**		**XpYp**
	ALLELE 1	ALLELE 2	ALLELE 1	ALLELE 2	
TTAGGG	0.11	0.070	0.069	0.13	0.071
TCAGGG	0.40	0.32	0.75	0.34	0.33
TTCGGG	0.91	0.81	0.95	0.92	0.83
GTAGGG	0.75	0.64	0.73	0.71	0.64
TGAGGG	0.33	0.20	0.26	0.66	0.24
TTGGGG	0.55	0.52	0.55	0.67	0.098
TAAGGG	0.23	0.096	0.34	0.47	0.082
CTAGGG	0.72	0.23	0.22	0.43	0.25
TTTGGG	0.57	0.077	0.25	0.26	0.047
TTAAGGG	0.22	0.082	0.33	0.50	0.073
TGCGGG	0.11	0.10	0.16	0.22	0.079
AGAGGG	0.026	0.022	0.32	0.049	0.031
OTHER	0.79	0.82	0.88	0.83	0.83
ALL	0.89	0.93	0.93	0.87	0.93

Table 5.2: Results from CLES tests for the HCT116$^{ATRX-/-}$ model.
Comparison of proportions of specific variants relative to the telomere length between the parental and the ALT clone at the 17p, 7q and XpYp chromosome ends. In black values <0.20; in green, values between 0.20 and 0.49; in orange, values between 0.50 and 0.79; and in red, values above 0.80.

5.4.3.2 HCA2 cell line

Similarly to the HCT116 cells, the variant repeat content for the parental and the ALT clone were plotted to highlight any differences (figure 5.9). Unlike the HCT116 cells, no distinct variant repeat proved to be more prevalent in the ALT clone across all chromosome ends and alleles. Although, most variant repeats showed an increase at the 17p chromosome end

whilst the XpYp chromosome end showed little change overall. This coincides with the STELA profiles at these chromosome ends with a larger range of telomeres in the ALT clone at the 17p telomeres (0.6-10.4 kb versus 1.1-6.1 kb in the parental; figure 5.2 C and D) compared to the XpYp telomeres (0.6-10.3 kb versus 0.5-9.6 kb in the parental) furthering the hypothesis made in chapter 3 that the 17p telomeres undergo larger insertion events as a result of ALT elongation.

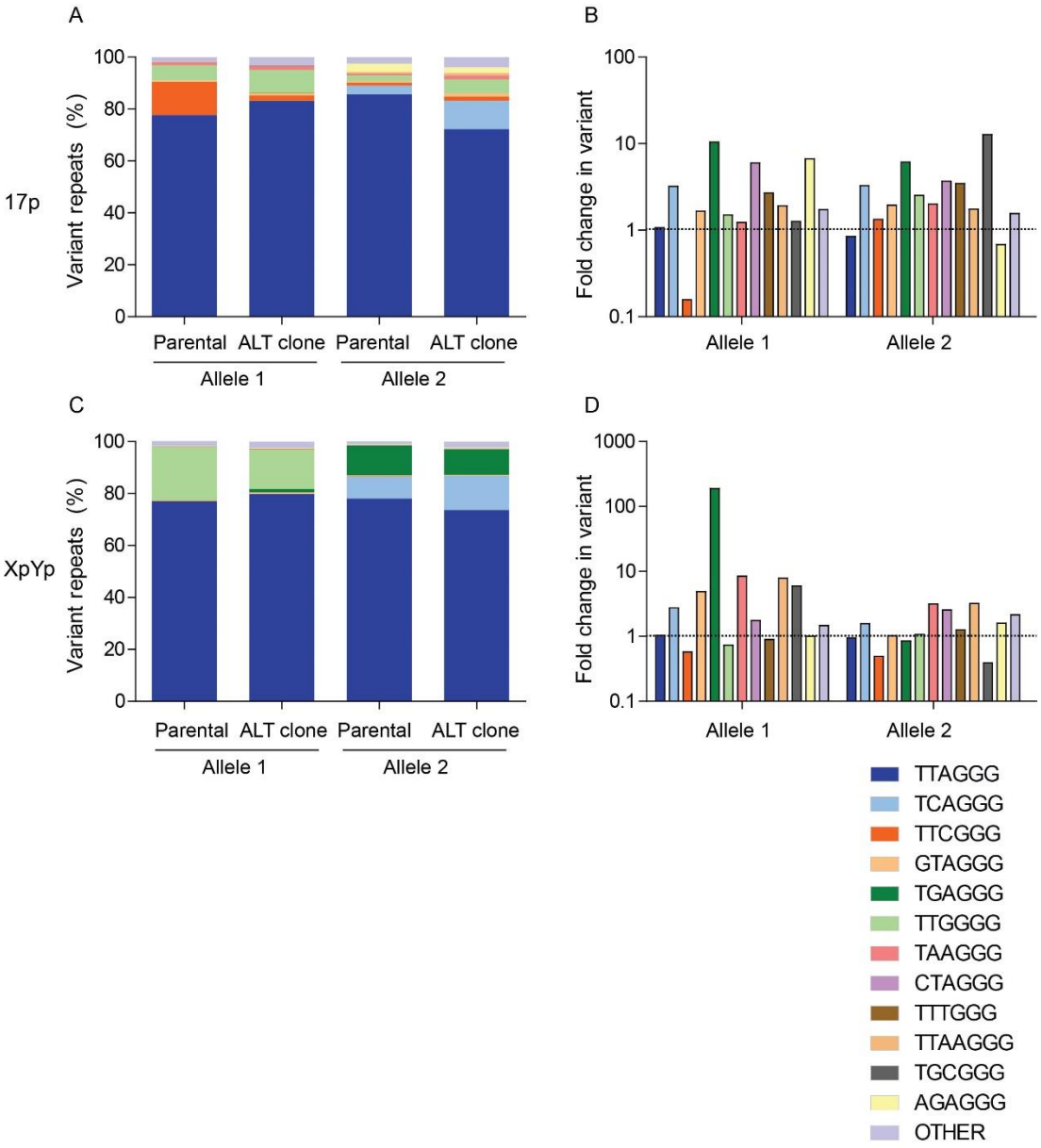


Figure 5.9: Telomere variant repeat proportions in HCA2$^{HPV\ E6E7}$ cells.
Total proportion of specific variant repeats combining all reads expressed in percentage for the parental and the ALT clone at the 17p (A); and XpYp (C) chromosome ends. Matching graph expressing the fold change in variant repeat proportion when comparing parental and ALT clone using a log scale for 17p (B); and XpYp (D).

Upon CLES testing, the TTCGGG repeat content, that was shown to be expanded in HCT116 cells, showed a no, to low effect size suggesting that the increase observed in this variant in the HCT116 model, maybe cell line specific. In addition, no specific variant repeat was consistently increased in the HCA2 model, although TCAGGG was increased in all but XpYp allele 1. Variant repeat content also appeared to be chromosome specific as seen with a medium effect size at the 17p chromosome end for the TAAGGG variant repeat, which showed no effect size at XpYp. Furthermore, the XpYp chromosome end showed little differences to the parental concerning variant repeat proportion overall. When comparing the overall variant repeat content, all telomeres showed an increased proportion, although only 17p allele 2 showed a large effect size, thus further supporting the hypothesis that the increase in non-canonical variant repeats is associated with the ALT mechanism.

	17p		**XpYp**	
	ALLELE 1	ALLELE 2	ALLELE 1	ALLELE 2
TTAGGG	0.60	0.12	0.60	0.39
TCAGGG	0.62	0.89	0.049	0.71
TTCGGG	0.043	0.39	0.044	0.037
GTAGGG	0.60	0.79	0.29	0.21
TGAGGG	0.079	0.060	0.11	0.42
TTGGGG	0.77	0.93	0.31	0.12
TAAGGG	0.56	0.71	0.15	0.10
CTAGGG	0.24	0.096	0.086	0.10
TTTGGG	0.11	0.077	0.11	0.020
TTAAGGG	0.12	0.58	0.13	0.097
TGCGGG	0.0063	0.0054	0.0066	0.0050
AGAGGG	0.013	0.18	0.39	0.48
OTHER	0.71	0.70	0.52	0.68
ALL	0.40	0.88	0.40	0.61

Table 5.3: Results from CLES tests for the HCA2$^{HPV\ E6E7}$ cell model.
Comparison of the proportions of specific variants relative to the telomere length between the parental and the ALT clone at the 17p and XpYp chromosome ends. In black values <0.20; in green, values between 0.20 and 0.49; in orange, values between 0.50 and 0.79; and in red, values above 0.80.

5.4.3.3 Comparison to U2OS

The telomere variant content of the U2OS cell line was plotted to highlight the incidence of specific variants (figure 5.10). Surprisingly, all chromosome ends studied showed a high proportion of the canonical telomere repeat TTAGGG, especially at the XpYp chromosome end with an overall occurrence of 96.8% (figure 5.10 B). Unlike the other models, especially

the HCT116 cells, there does not appear to be an increased proportion of non-canonical variant repeats; although allele 1 at the 17p chromosome end, appears to have a higher rate of variants thus potentially suggesting that ALT may be allele specific.

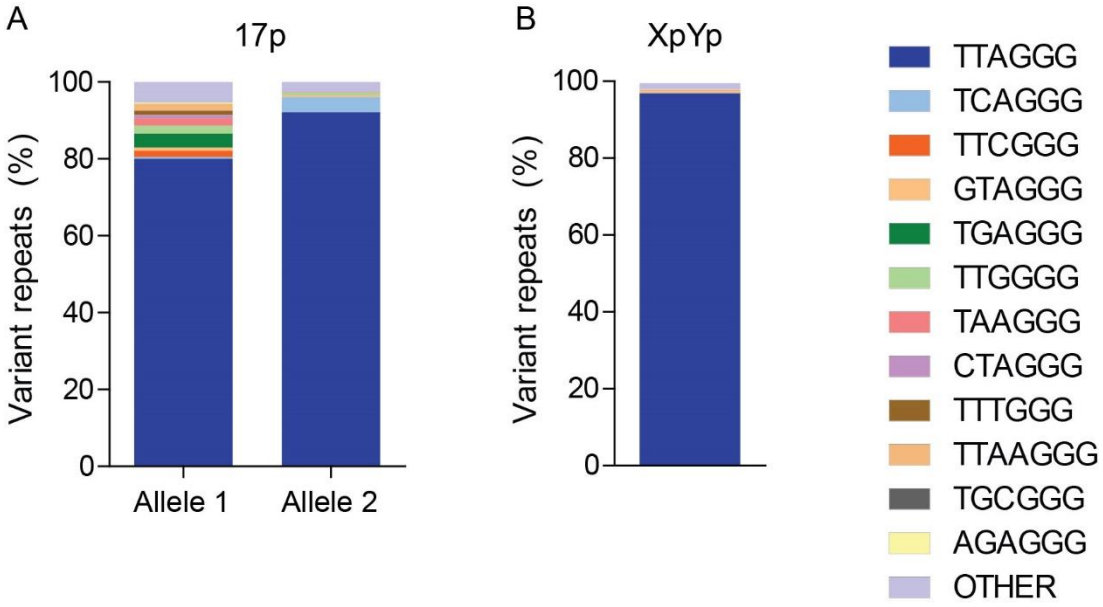


Figure 5.10: Telomere variant repeat proportion in U2OS cells.
Total proportion of specific variant repeats combining all reads expressed in percentage for the U2OS cell line at the 17p (A); and XpYp (B) chromosome ends.

5.4.4 Characterisation of ALT elongation events

5.4.4.1 ALT events are dependent on the cell type

To distinguish patterns in elongation events within ALT telomeres, compared to the parental cells, single telomeres were selected to highlight different variant repeat patterns. Variants were colour-coded and condensed to highlight patterns of ALT-elongation (figure 5.11). The increase in the TTCGGG variant repeat in the HCT116 model was highlighted at all chromosome ends and alleles, thus suggesting that ALT affects all alleles irrespective of telomere length. Different patterns of TTCGGG variant enrichment were observed with the interspersion among TTAGGG repeats as well as large blocks constituting up to 82% of the whole telomere sequence.

In contrast, the HCA2 cells appeared to have an increase in TTAGGG repeats as the main elongation event. The reduction in TTCGGG repeats was further confirmed as a result of telomere erosion beyond the point where the series of TTCGGG-rich regions appear in allele 1 at the 17p telomeres. Furthermore, the XpYp allele 1 showed an increase in TGAGGG

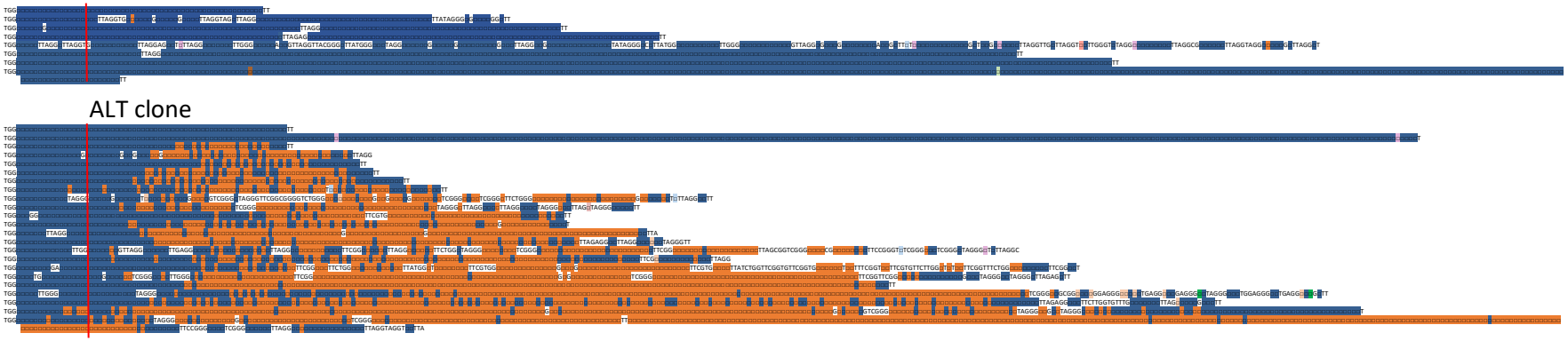
repeats which are observed in the parental allele 2 but not in allele 1 thus perhaps suggesting an inter-chromosomal exchange, further explaining the enrichment of the TTAGGG repeats in allele 2 thus using allele 1 as a template.

The U2OS telomeres further confirmed the high proportion of TTAGGG repeats which further suggests that ALT is cell type dependent. Furthermore, the 17p chromosome end allele 1 showed a higher proportion of non-canonical variant repeats in figure 5.10. Upon analysis of the telomere sequence, these variant repeats appeared at the start of the telomere and are suggestive of natural allelic variability within the first 1 kb of the telomere repeat array, rather than as a consequence of ALT elongation. These data indicate that the overall increase in variant repeats observed in the HCA2$^{HPV\ E6E7}$ and especially HCT116$^{ATRX-/-}$ model can result from ALT elongation however, the nature of the repeat sequences utilised appears to be cell line specific.

HCT116 cells

XpYp telomere

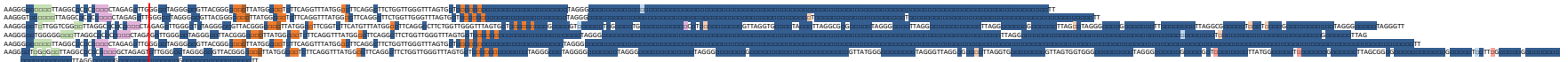
Parental



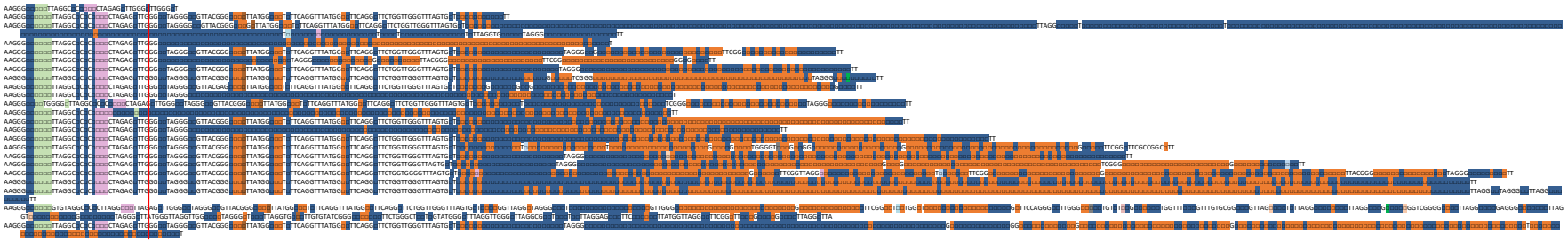
ALT clone

17p telomere Allele 1

Parental



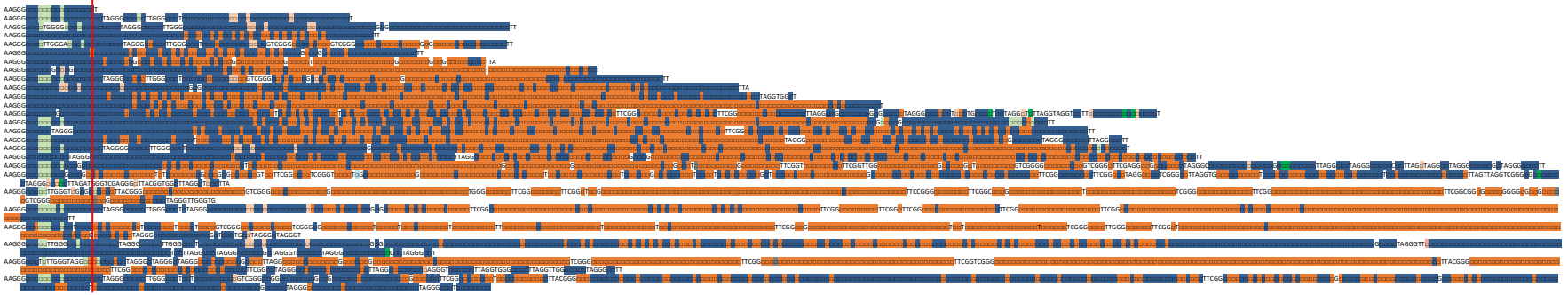
ALT clone



17p telomere Allele 2

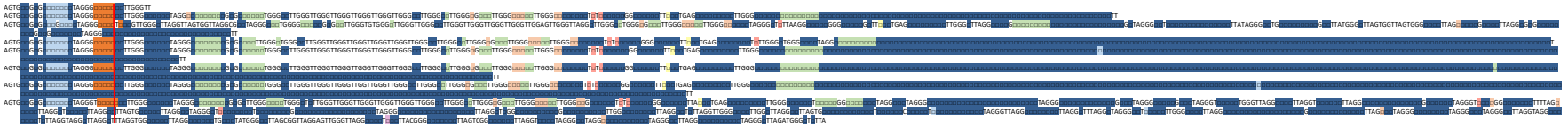
Parental

ALT clone

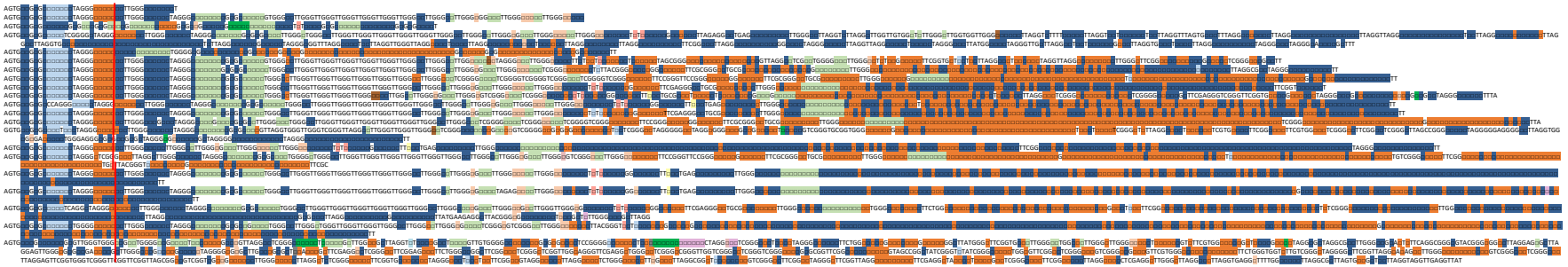


7q telomere Allele 1

Parental

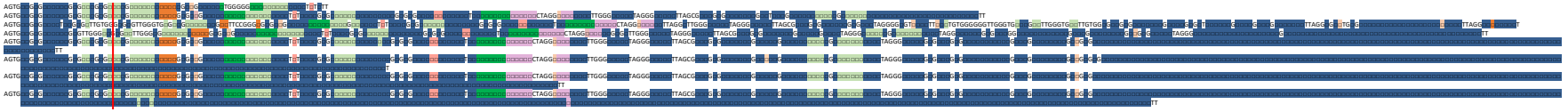


ALT clone

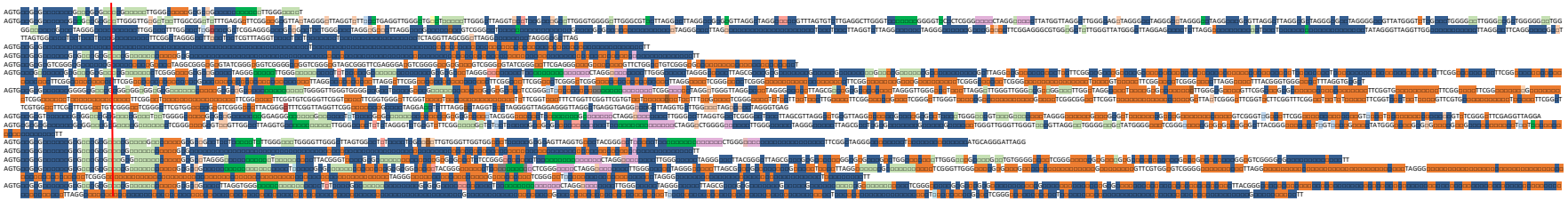


7q telomere allele 2

Parental



ALT clone

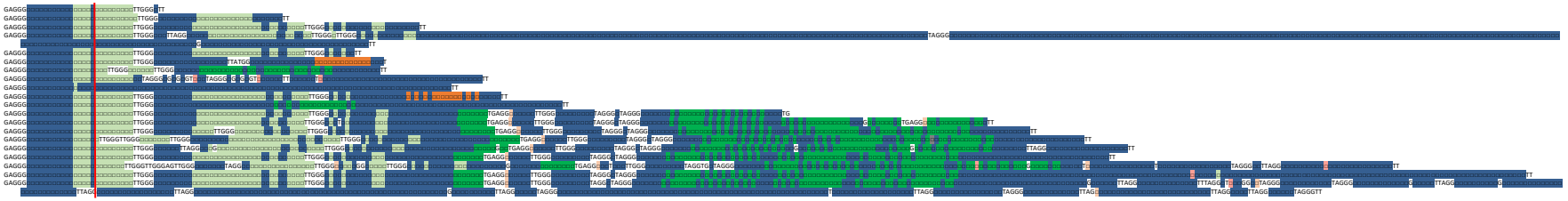


HCA2 cells

XpYp telomere allele 1

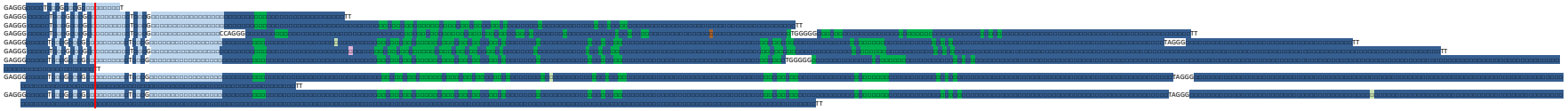
Parental

ALT clone

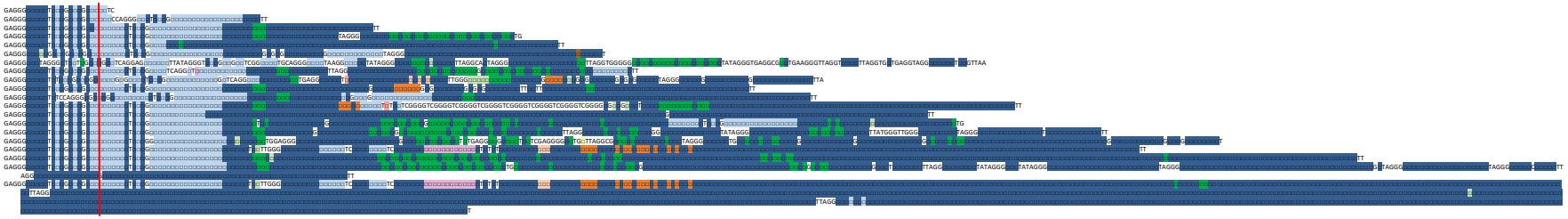


XpYp telomere allele 2

Parental

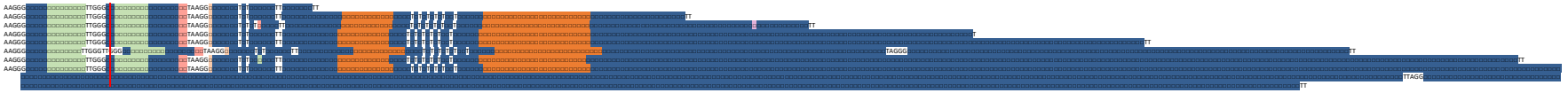


ALT clone

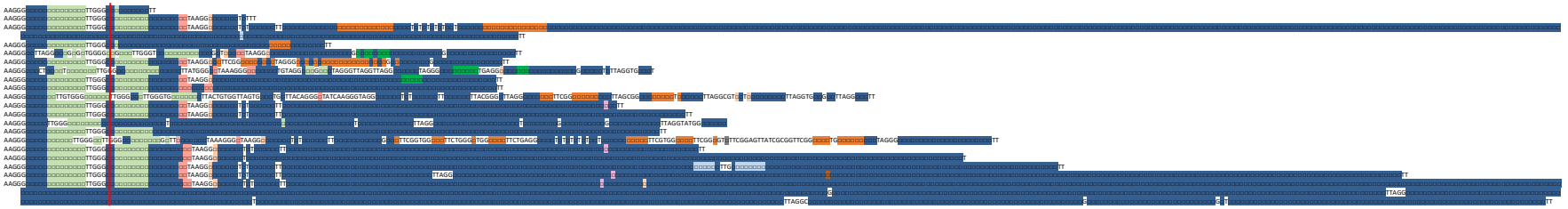


17p telomere allele 1

Parental

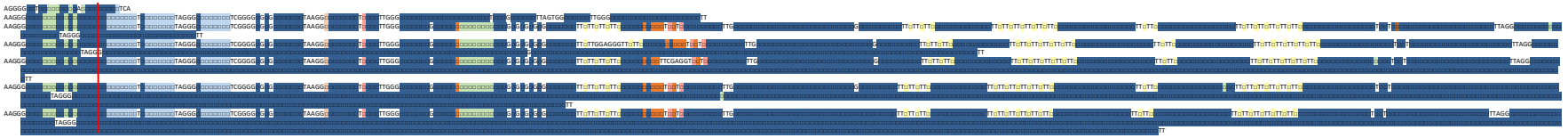


ALT clone

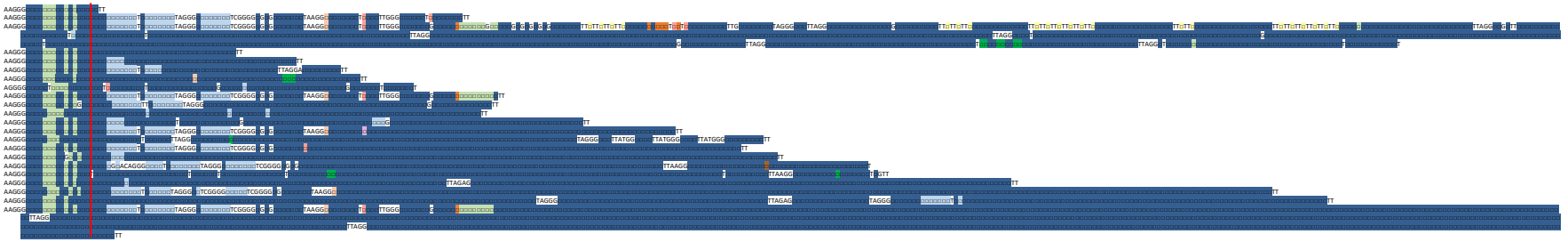


17p telomere allele 2

Parental

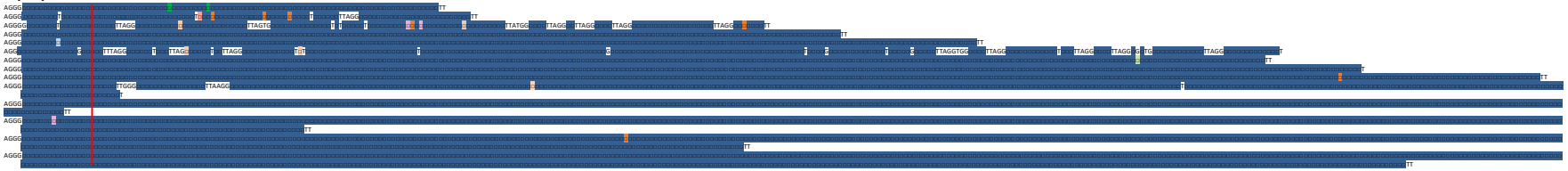


ALT clone

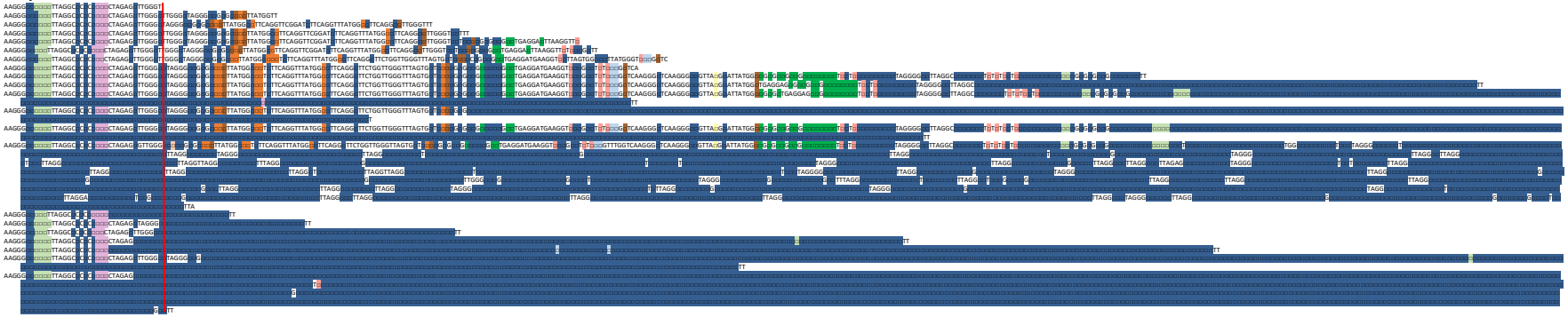


U2OS cells

XpYp telomere



17p telomere allele 1



17p telomere allele 2

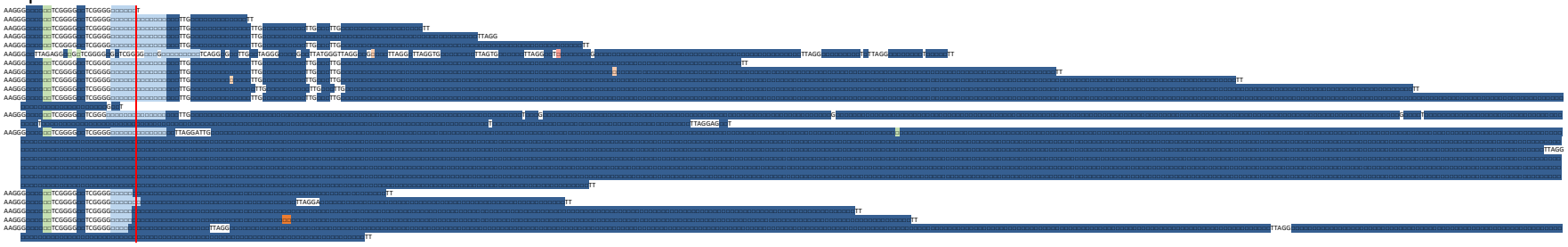


Figure 5.11: Telomere repeat content alteration in ALT cells.

Telomere sequences of A) HCT116$^{ATRX-/-}$ cells (parental followed by ALT clone) at the XpYp, 17p alleles 1 and 2 and 7q alleles 1 and 2; B) HCA2$^{HPV\ E6E7}$ cells (parental followed by ALT clone) at the XpYp alleles 1 and 2 and 17p alleles 1 and 2; and C) U2OS cells at the XpYp and 17p alleles 1 and 2. Legend as follows: TTAGGG; TCAGGG; TTCGGG; GTAGGG; TGAGGG; TTGGGG; TAAGGG; CTAGGG; TTTGGG; TGCGGG; AGAGGG. □ represents a 6 nucleotide variant repeat and non-highlighted sequences represent single nucleotides. The red line indicates the end of the first 100 bps used for differentiation of alleles.

5.4.4.2 ALT elongation is a result of multiple events

To attempt to understand the mechanism of elongation and establish the occurrence of a pattern, a selection of reads were analysed to assess if all alleles undergo elongation and if this is a result of a single or multiple events. The HCT116 17p and the HCA2 XpYp chromosome ends are presented in figure 5.12 whilst other examples are presented in appendix 4. It was previously hypothesised (chapter 4), that at the 17p chromosome end in the HCT116 model, the shorter allele undergoes elongation to a mean of 6.68 kb, whilst the longer allele prior to crisis simply erodes. This would suggest that if the enrichment in TTCGGG repeats is distinctive of ALT elongation, only one allele would have that distinct increase in this variant. However, the TTCGGG repeat patterns were observed at both alleles therefore suggesting that all alleles undergo elongation irrespective of telomere length. In the examples shown in figure 5.12, each telomeric sequence displayed a different extent of telomere erosion and replacement with the TTCGGG repeat interspersion patterns, an estimated length that the telomere has been reduced to is annotated in red (figure 5.12). These data indicated that the length at which telomeres are subjected to TTCGGG repeat replacement is variable. Moreover, as each telomeric sequence is distinct both in terms of the repeat variant interspersion pattern and the insertion point, this indicates that ALT elongation arises from multiple events and is therefore not clonal. In addition, telomeres from allele 2 appeared to erode to shorter lengths than allele 1 (as short as 79 bps at allele 2 versus 104 bps at allele 1) which could imply that allele 2 was the shortest prior to crisis. Due to an issue with loading of this sample at the time of sequencing, limited reads over 5 kb were obtained therefore limiting the conclusions that can be drawn from this experiment.

Upon analysis of HCA2 XpYp telomeres, similar patterns of erosion and enrichment of specific variant repeats, such as TGAGGG or TTAGGG, also support the concept that ALT elongation is a result of multiple events (figure 5.12) and that telomeres erode to different lengths before a change in variant pattern sequence is observed. In addition, the examples below highlight the increase in TGAGGG at allele 1 and TTAGGG at allele 2 at the XpYp chromosome end further supporting the hypothesis that ALT elongation could be a result of inter-chromosomal HDR.

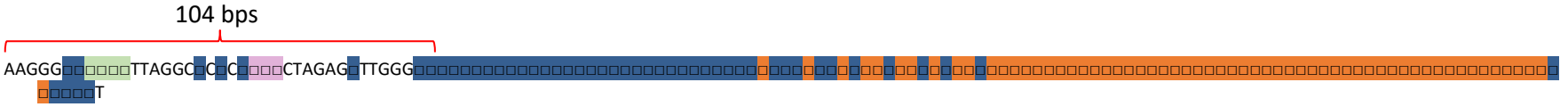
HCT116$^{ATRX-/-}$

17p allele 1

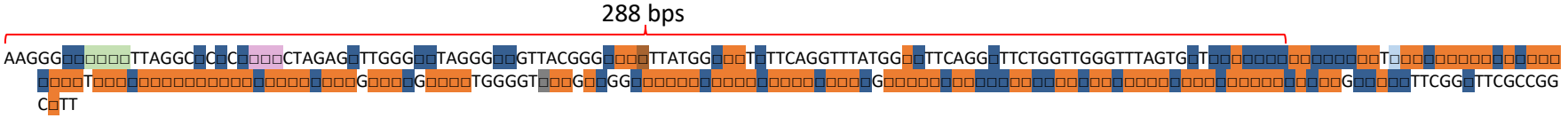
- Parental (reference)

- ALT clone

104 bps



288 bps

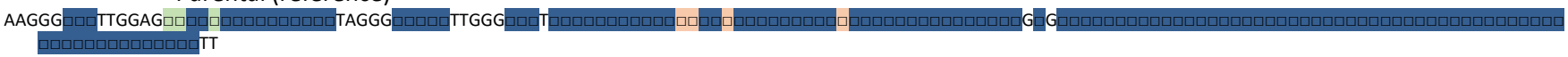


342 bps

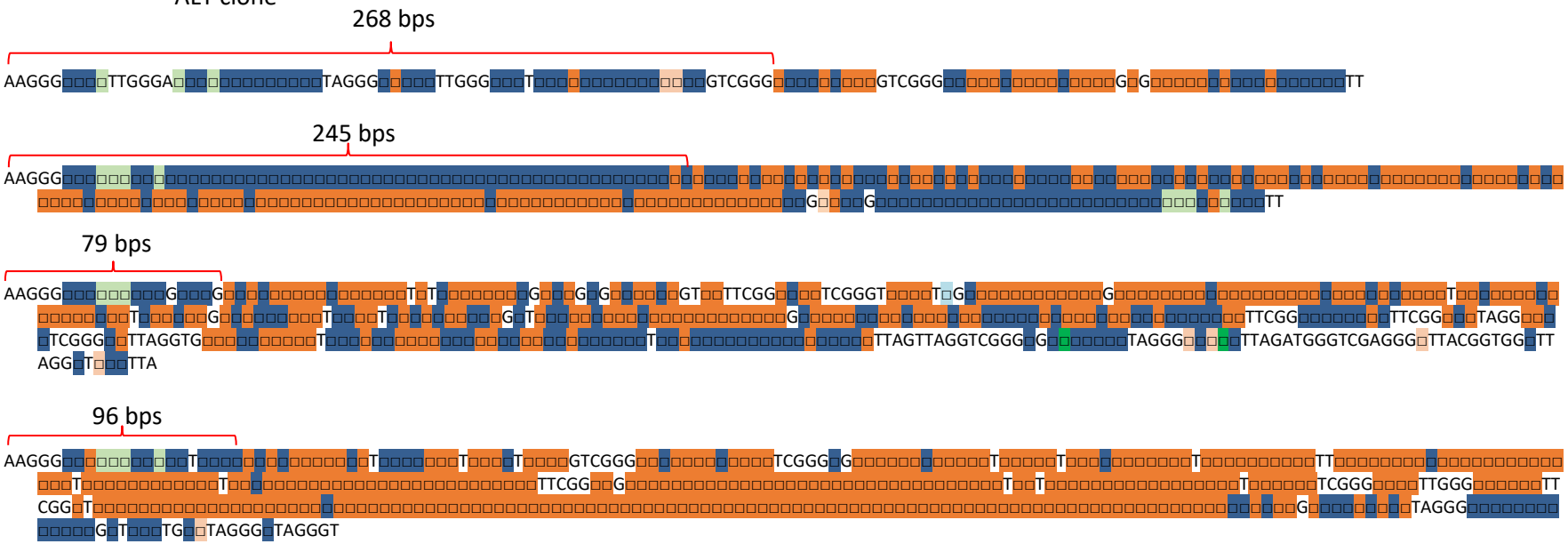
992 bps

17p allele 2

- Parental (reference)



- ALT clone



HCA2$^{HPV\ E6E7}$

XpYp allele 1

- Parental (reference)

- ALT clone

(unaltered)

201 bps

178 bps

214 bps

536 bps

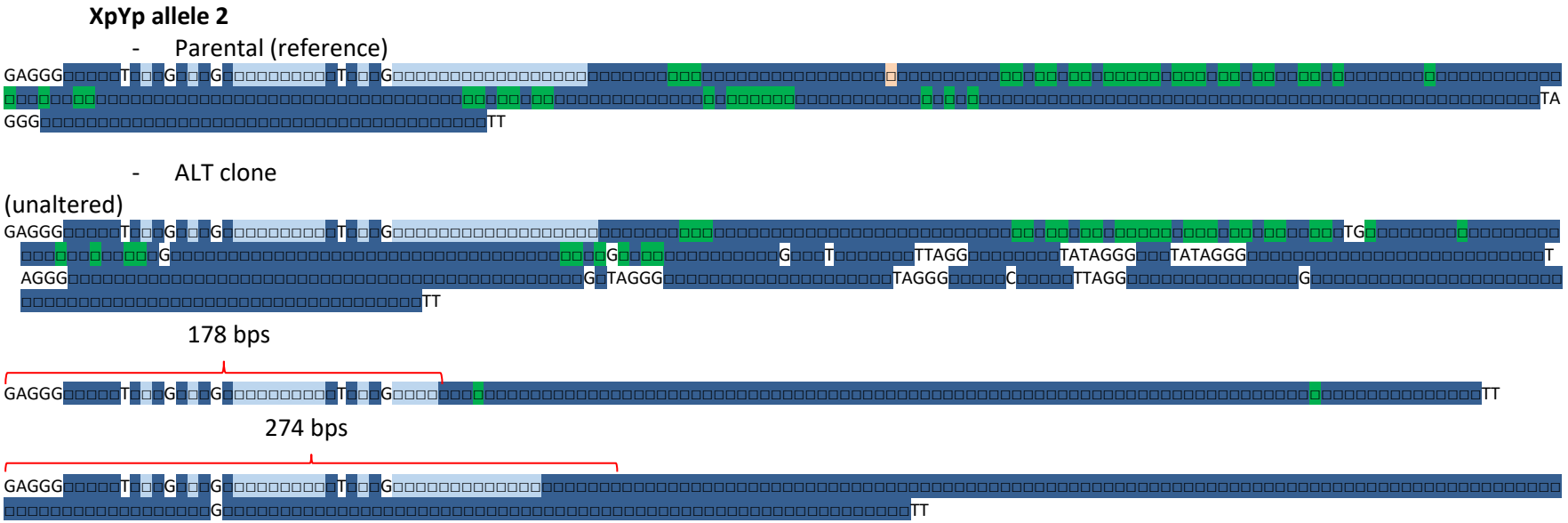


Figure 5.12: ALT elongation results from multiple events.
Examples of ALT telomere sequences in the HCT116$^{ATRX-/-}$ parental and ALT clone at the 17p chromosome end at both alleles; and in the HCA2$^{HPV\ E6E7}$ parental and ALT clone at the XpYp chromosome end at both alleles. Estimated length to which telomeres eroded to prior to change in sequence as a result of ALT detailed for each read by red line.

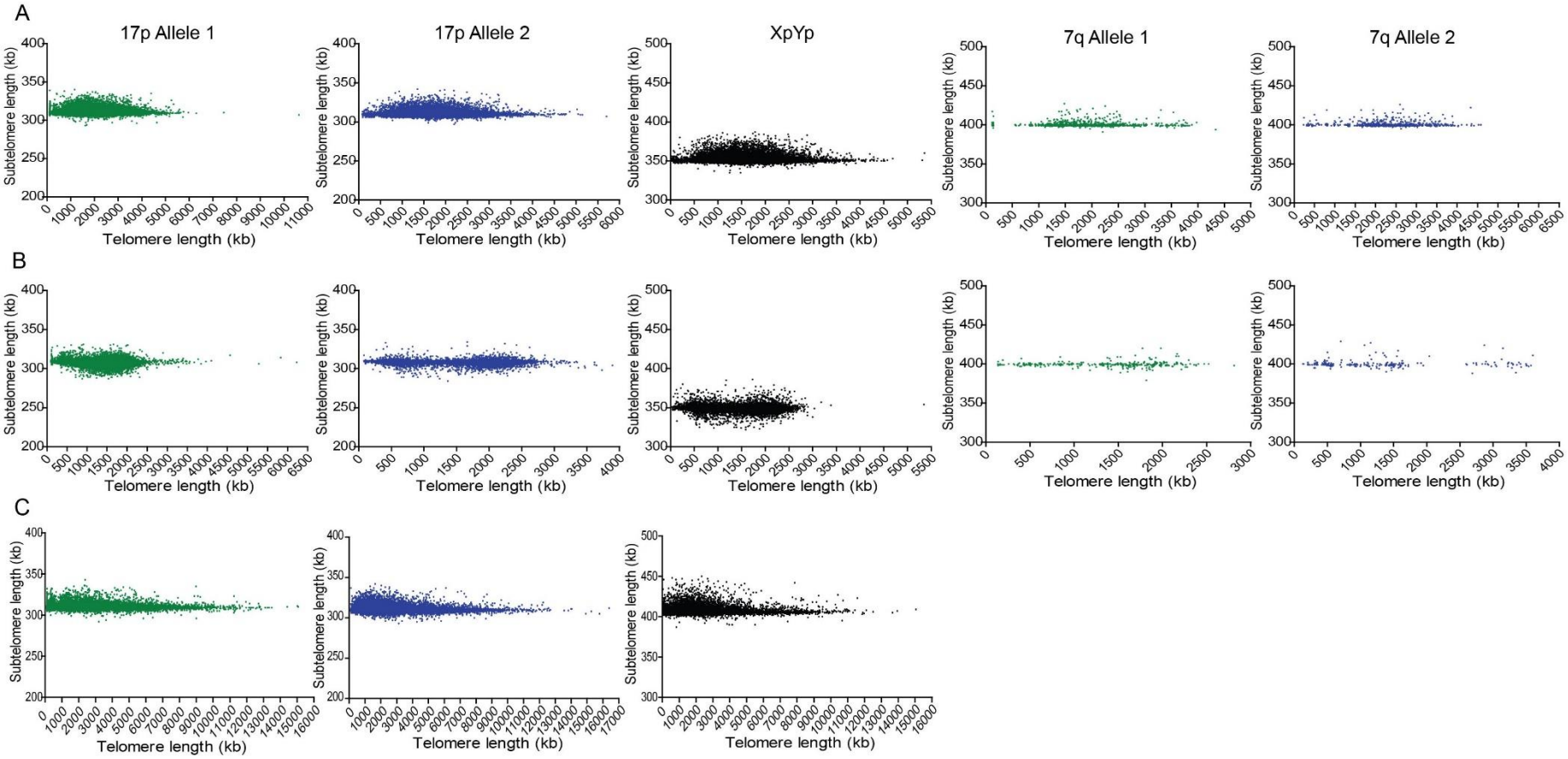
5.4.5 Sub-telomere region alterations

The data presented above provided evidence that telomeres can erode to within 79 bps of the beginning of telomere repeat array and that ALT appears to be the result of multiple independent events (figure 5.12). It was therefore considered that telomere erosion may continue into the telomere-adjacent DNA, and this may be lead to deletion events into the sub-telomere DNA, followed by the synthesis of new telomere variant repeats associated with crisis or perhaps ALT upregulation.

Sub-telomere regions that matched the primer used to generate the STELA profiles were analysed. As previously mentioned, the start and end of the sub-telomere and telomere sequences were determined by using the Hidden Markov Model. By this means, the start of the telomere is subject to starting with variant repeats that are recognised by this model, with a maximum error rate of 2 bps from the canonical repeat, and minor changes can alter the start of the trimmed telomere region. Small insertions and deletions at the sub-telomere were therefore anticipated for this reason. In addition, sequencing is associated with a small error rate and therefore any sub-telomere regions within 10% or the expected sub-telomere length were not considered to be significant. Most reads were within the expected length (99.8 – 100%) (figure 5.13) however, any sub-telomere sequences with an insertion or deletion over the length stated in table 5.4, were analysed for potential unique recombination events alongside the telomere attached. These consistently mapped to the appropriate sub-telomere sequence despite insertions and deletions suggestive of sequencing errors associated with PacBio sequencing (appendix 5).

	XpYp (bps)	**17p (bps)**	**7q (bps)**
HCT116 cells	35	30	40
HCA2 cells	41	300	N/A
U2OS cells	41	30	N/A

Table 5.4: 10% of total sub-telomere length for each chromosome end studied.
The number of base pairs that represent 10% of the expected sub-telomere length are stated for each chromosome end in each sample and used to account for natural error rate during the sequencing process.


A
17p Allele 1
17p Allele 2
XpYp
7q Allele 1
7q Allele 2
B
C
Subtelomere length (kb)
Telomere length (kb)

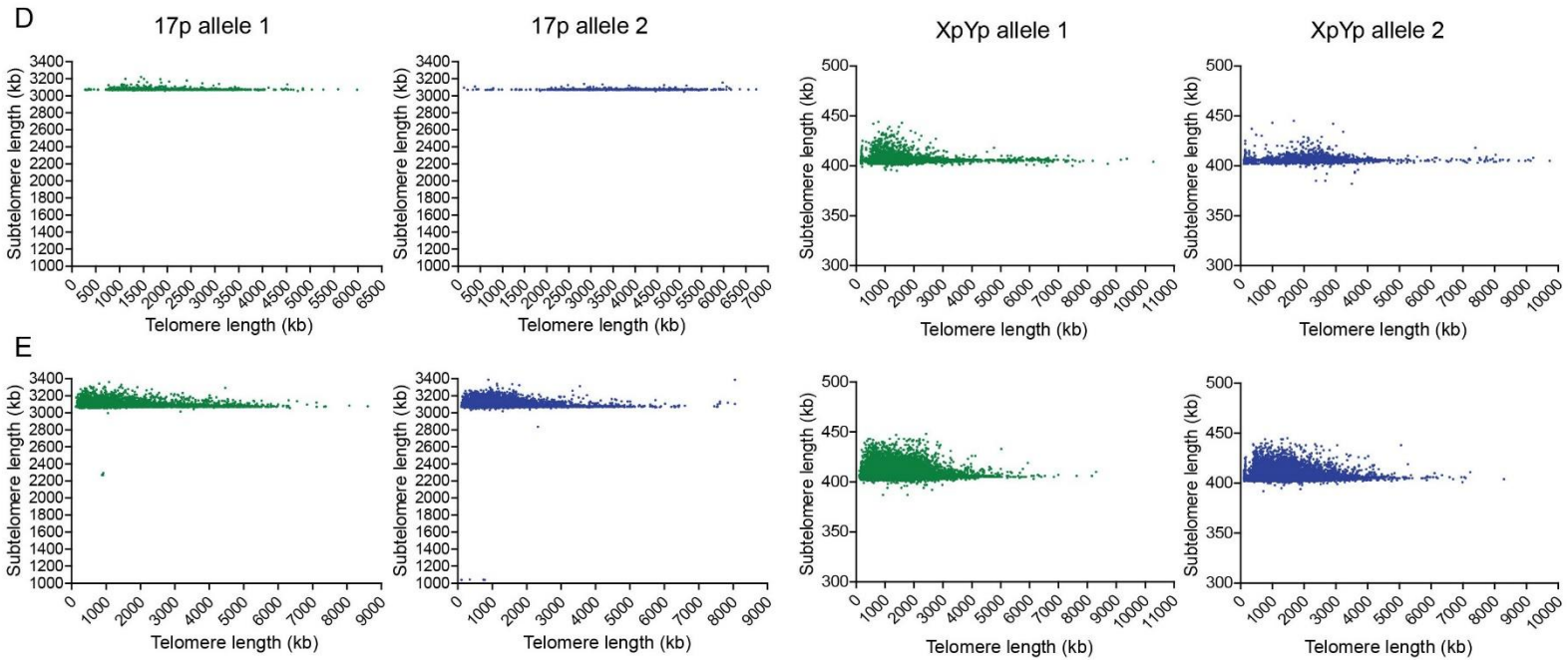


Figure 5.13: Reads with sub-telomere DNA sequence matching the telomere specific primer have limited alteration to the sub-telomere sequence.
Graphs plotting the sub-telomere length in kb (y-axis) against the telomere length in kb (x-axis) of A) HCT116$^{ATRX-/-}$ parental; B) HCT116$^{ATRX-/-\ DN-hTERT}$ ALT; C) U2OS; D) HCA2$^{HPV\ E6E7}$ parental; and E) HCA2$^{HPV\ E6E7\ ATRX-/-}$ ALT at the 17p (left), XpYp (middle) and 7q (right) chromosome ends.

5.4.6 Insertions

To assess if specific DNA sequences were inserted into the telomere as a result of elongation through ALT, insertions were isolated and analysed. Two types of insertions were observed: interstitial insertions which occurred within the telomere repeat array and included in the overall telomere length; and end insertions which occurred after the telomere repeat array and before the telorette primer sequence and not included in the overall telomere length (refer to figure 5.5). These insertions differ from the TTAGGG repeat by over of 2 bps thus, variant repeats are not included in background sequence. The proportion of both interstitial and end insertions appeared to be increased in ALT clones when comparing to the respective parental cell lines (figure 5.14 A and B). In addition, the HCT116$^{ATRX-/-}$ parental had more insertions than the HCA2$^{HPV\ E6E7}$ parental indicating an increased background of insertions potentially originating from the crisis phase these cells have undergone before becoming cancerous. The HCT116$^{ATRX-/-}$ ALT clone presented the most insertions and was consistently significantly different from the parental following a Fisher's exact test, apart from end insertions at the 7q allele 2 (table 5.6). A statistically significant difference was only seen at the XpYp chromosome end in the HCA2 cell line, therefore, the increased rate of insertions in the HCT116$^{ATRX-/-}$ ALT clone was not attributed to the ALT mechanism but rather as a consequence of crisis.

Samples compared	**Telomere**	**Fisher test result**	
		Interstitial	End
HCT116$^{ATRX-/-}$ parental vs ALT clone	17p allele 1	< 0.0001	< 0.0001
	17p allele 2	< 0.0001	< 0.0001
	XpYp	< 0.0001	< 0.0001
	7q allele 1	< 0.0001	0.0069
	7q allele 2	< 0.0001	0.070
HCA2$^{HPV\ E6E7}$ parental vs ALT clone	17p allele 1	0.97	0.92
	17p allele 2	0.98	0.99
	XpYp allele 1	< 0.0001	< 0.0001
	XpYp allele 2	0.0002	0.55

Table 5.5: Summary of the Fisher's exact test results comparing the insertion rates between controls and ALT clones.
Parentals (HCT116$^{ATRX-/-}$ and HCA2$^{HPV\ E6E7}$) compared to their respective ALT clone (HCT116$^{ATRX-/-\ DN-hTERT}$ clone 2 and HCA2$^{HPV\ E6E7\ ATRX-/-}$ clone 21) at the chromosome ends studied. Statistical differences highlighted in red ($P < 0.05$).

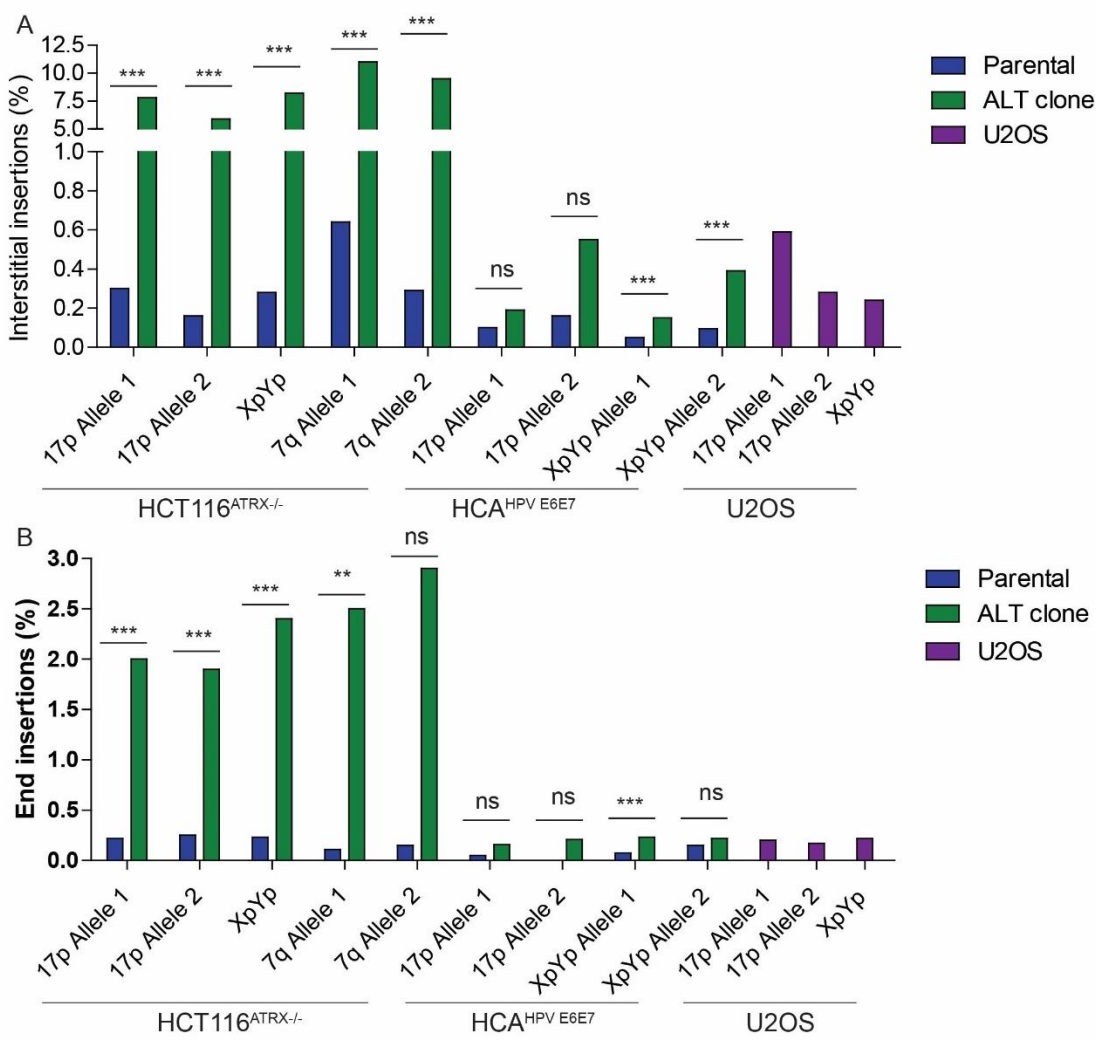


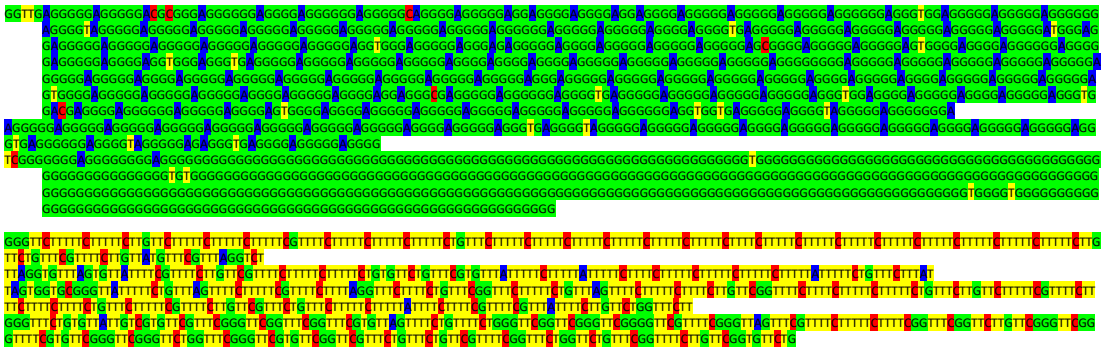
Figure 5.14: Proportion of insertions within telomere sequences.
A) Proportion of reads with at least one interstitial insertion. B) Proportion of reads with at least one end insertion. In blue the parentals (HCT116$^{ATRX-/-}$ and HCA2$^{HPV\ E6E7}$); in green the ALT clones (HCT116$^{ATRX-/-\ DN-hTERT}$ clone 2 and HCA2$^{HPV\ E6E7\ ATRX-/-}$ clone 21); and in purple the U2OS cell line.

Analysis of the insertion nucleotide sequence, for both the interstitial and end insertions, presented G- or T-rich sequences that were therefore not mappable to specific loci in the genome (figure 5.15). This was observed across all samples. In addition, the HCT116 cells showed a background of sub-telomere insertions within the telomere repeat array, such as 17p sub-telomere sequence within XpYp telomeres, and these were therefore considered to be associated with PCR-related recombination events as both the XpYp and 17p telomeres were amplified together (figure 5.15). In addition, due to the occurrence of these repetitive insertions across all samples irrespective of ALT initiation or cells undergoing crisis, these were not associated with the ALT mechanism and could be an artefact originating from the PacBio sequencing platform.

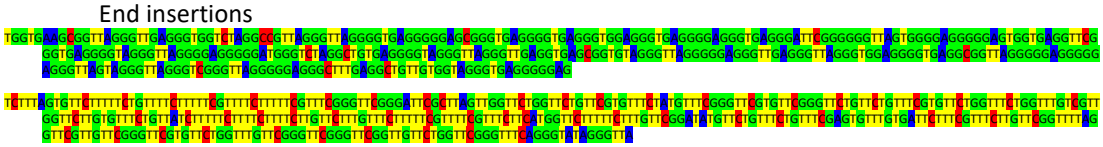
XpYp chromosome end

HCT116$^{ATRX-/-\ DN\text{-}hTERT}$ ALT clones

Interstitial insertions

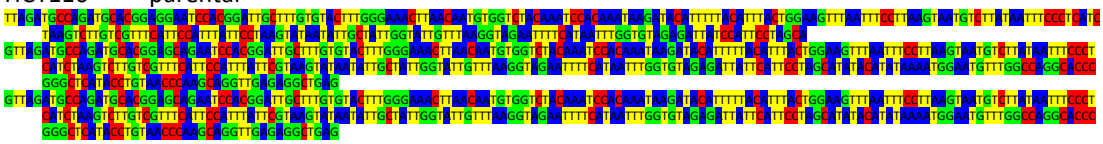


End insertions



17p sub-telomere sequence within XpYp telomere repeat array

HCT116$^{ATRX-/-}$ parental



HCT116$^{ATRX-/-\ DN\text{-}hTERT}$ ALT Clone

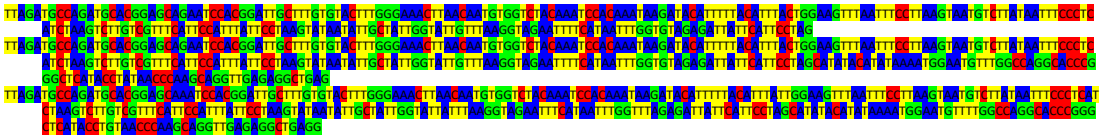


Figure 5.15: Examples of interstitial and end insertions commonly found in all samples.
Insertions from the HCT116$^{ATRX-/-}$ parental and ALT clone at the XpYp chromosome end with guanine in green, adenine in blue, thymine in yellow and cytosine in red.

5.4.7 Impact on the genome

Cells that undergo a telomere-driven crisis and escape are subject to chromosomal rearrangements which drive cancer genome evolution. The impact of crisis in the context of ALT is not readily studied in the literature. Therefore, to assess the impact the loss of ATRX combined with a telomere-driven crisis have on the genome, a range of HCT116$^{ATRX-/-\ DN\text{-}hTERT}$ samples were sent for WGS. Samples were picked to represent three categories (ALT-survived; ALT-died; and Telomerase positive). Where possible, pre- and post-crisis samples were sequenced to assess if specific alterations occur when ALT is upregulated. These samples are summarised in table 5.6 and were sequenced using the BGISEQ-500 platform providing a paired-end reads with a 15x coverage. The HCT116 WT structural variants (SV)

and copy number variants (CNV) were classed as background and subtracted from each clone. The data were mapped and analysed by Dr Kez Cleal.

Category	Sample	In relation to crisis	Characteristics
ALT-survived	Cl2 p2	Pre-	
	Cl2 p5	Post-	ALT no telomerase
	Cl3 p2	Pre-	
	Cl3 p5	Post-	ALT no telomerase
	Cl4 p2	Pre-	
	Cl4 p5	Post-	ALT no telomerase
	Cl48 p5	Post-	ALT + telomerase
	Cl131 p2	Pre-	ALT no telomerase
	Cl131 p7	Post-	Telomerase
	Cl147 p2	Pre-	ALT no telomerase
	Cl147 p5	Post-	Telomerase
ALT-died	Cl37 p3	Pre-	ALT no telomerase
	Cl51 p2	Pre-	ALT no telomerase
	Cl52 p4	Pre-	ALT no telomerase
	Cl81p p2	Pre-	ALT no telomerase
	Cl83 p2	Pre-	ALT no telomerase
	Cl100 p2	Pre-	ALT no telomerase
	Cl108 p2	Pre-	ALT no telomerase
	Cl130 p5	Pre-	ALT no telomerase
	Cl132 p3	Pre-	ALT no telomerase
	Cl148 p5	Pre-	ALT no telomerase
Telomerase +	Cl49 p2	Pre-	
	Cl49 p5	Post-	Telomerase
	Cl50 p5	Post-	Telomerase
	Cl53 p2	Pre-	
	Cl53 p9	Post-	Telomerase
	Cl92 p4	Post-	Telomerase
	Cl103 p4	Post-	Telomerase
	Cl111 p6	Post-	Telomerase
	Cl112 p5	Post-	Telomerase

Table 5.6: Summary of clones sent for WGS.
Details of the sampling point, relativity to crisis and the characteristic of each sample.

5.4.7.1 Structural and copy number variants

Structural variants (SV) include breakpoints, translocations, duplications, inversions and deletions and were called using TIDDIT and plotted for each clone (figure 5.16 A) by Dr Kez Cleal. No SVs were noted prior to crisis in all clones irrespective of the survival outcome and the mechanism upregulated for prolonged proliferation. Strikingly, SVs were only observed when telomerase was upregulated, with a complete lack of SVs in ALT-died clones, pre-crisis samples and ALT clones prior to the upregulation of telomerase (clone 2, 3 and 4 p5). This suggests that escape from crisis by upregulating ALT does not immediately incur SVs, as opposed to cells that upregulate telomerase. Furthering this observation, SVs were observed in ALT-survived clones 48 p2; 131 p7 and 147 p5 but only at the point at which telomerase was upregulated. A statistical difference was observed between the ALT-survived and telomerase positive clones when compared to the ALT-died clones (P value of 0.045 and 0.0017 respectively following Mann-Whitney test) whilst no difference was noted between the ALT-survived clones and telomerase positive clones (P value of 0.32).

Next, unique copy number variants (CNVs) were counted according to a previously described method (Cleal et al. 2019). Copy number gains were commonly observed although only statistically increased in the telomerase positive samples (p value = 0.029) whilst the overall number of CNVs was increased in the ALT-survived category (figure 5.16 B and C). One sample, (clone 147) appeared to be chromothriptic and remained in crisis for a long period (100 days). This explains the high standard deviation error bar regarding the ALT-survived category. No sign of chromothripsis was observed in other clones and this phenomenon was not associated with ALT, although it cannot be excluded that ALT activity had an impact on genome stability. Finally, a common amplification of chromosome 5, notably the p arm, was observed in 10 clones whilst the rest of the genome remained stable (figure 5.16 D).

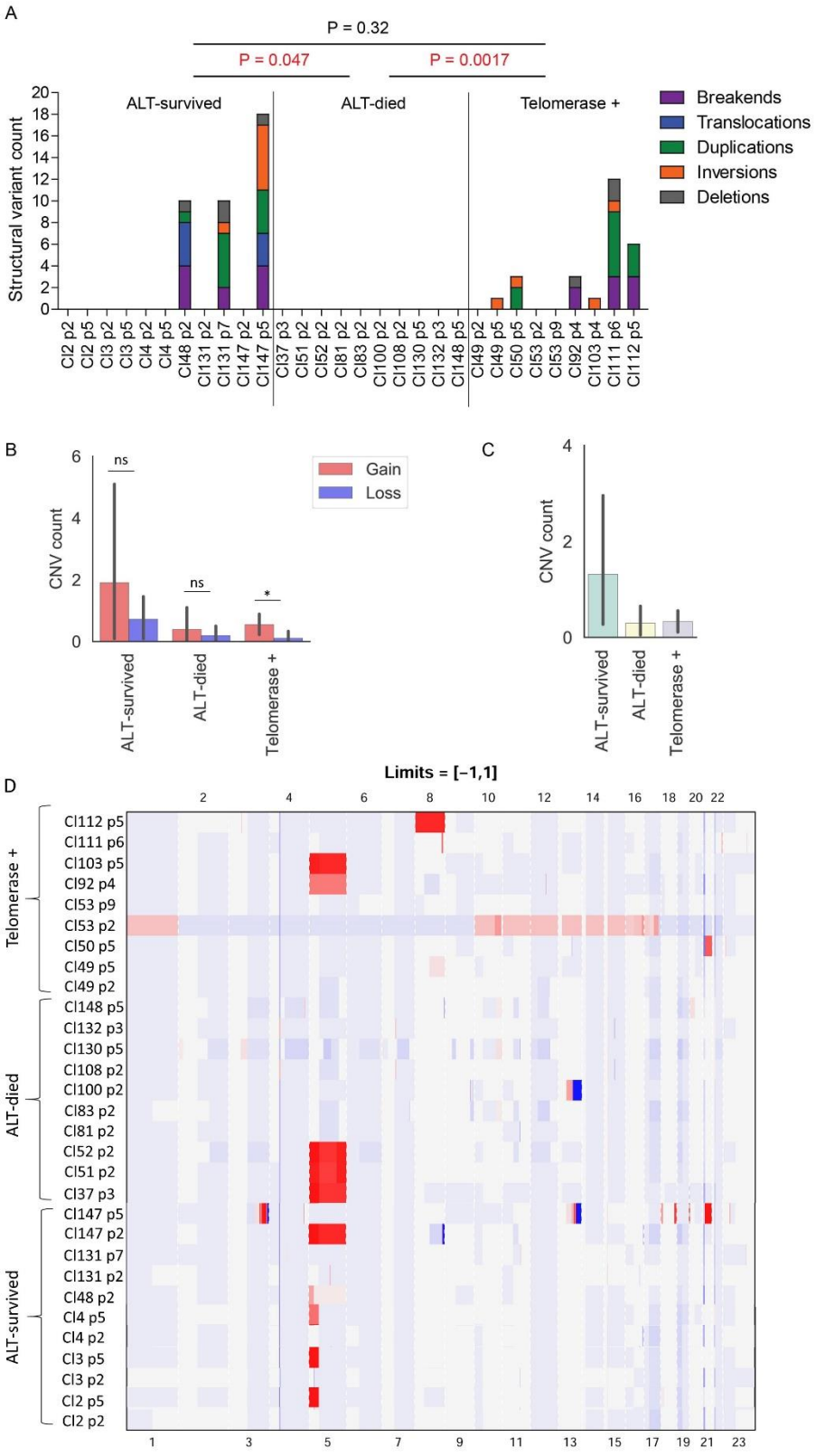


Figure 5.16: Structural and copy number variants following WGS of HCT116$^{ATRX-/-\ DN-hTERT}$ samples.
A) Structural variant counts for ALT-survived, ALT-died and telomerase positive clones with the P value stated above as a result of a Mann-Whitney test. Statistical difference highlighted in red (P value <0.05). Copy number variant counts showing loss and gains (B) and overall change (C). D) Heatmap of locations of copy number changes in each clone with the chromosome number listed across the top (even) and the bottom (uneven). With red signifying gains and blue signifying losses.

5.4.7.2 Clonality of clones

Despite isolating single cell clones, the potential for polyclonal populations is present and the clonality of all clones was therefore assessed for further analysis. Clonality was established using the variant allele frequency (VAF) method combined with the copy number (CN) method by Dr Kez Cleal, as previously described (Cleal et al. 2019). Firstly, the VAF method takes into account unique single nucleotide polymorphisms (SNPs) and a VAF of 0.5 signifies the presence of one heterozygous clone (two different alleles; figure 5.17 A and B) whilst a VAF of 0.25 is indicative of a polyclonal population (four different alleles; figure 5.17 C). However, if two clonal populations are not sufficiently diverse, this method will not highlight any differences and a larger number of variants must be analysed through the CN method which looks at all copy number changes. By this means, ideal peaks, indicative of a monoclonal population, would have an allele frequency of 0.5 at CN 0 and, if any gains are observed, a triploid state will be observed at CN+1 with two peaks at 0.33 and 0.66 frequencies (figure 5.17 B). The presence of another clone will alter these peaks (figure 5.17 D).

The status of each clone is presented in table 5.7 alongside the length of crisis to assess if there is any correlation between the two. Clones that did not undergo a visible crisis phase, such as clone 53 and 112, could be polyclonal and therefore show similarities with the MP which proliferated and immortalised without a noticeable crisis phase. By this means, it was confirmed that clone 53 was polyclonal. In addition, clone 48, which immortalised by upregulating telomerase and without undergoing telomere elongation despite being positive for C-circles, showed a co-existence of both ALT and telomerase on the long-term. This clone was confirmed to be polyclonal and suggests the co-existence of two populations with different telomere maintenance mechanisms rather than two mechanisms co-existing within one population. Interestingly, clones 4 and 147 were monoclonal prior to crisis but became polyclonal after crisis which could once more suggest the co-existence of two populations using different TMMs. A remaining 3 clones (100, 108 and 111) proved to be polyclonal, despite no indication on the STELA blots, whilst the remaining clones were confirmed to be monoclonal.

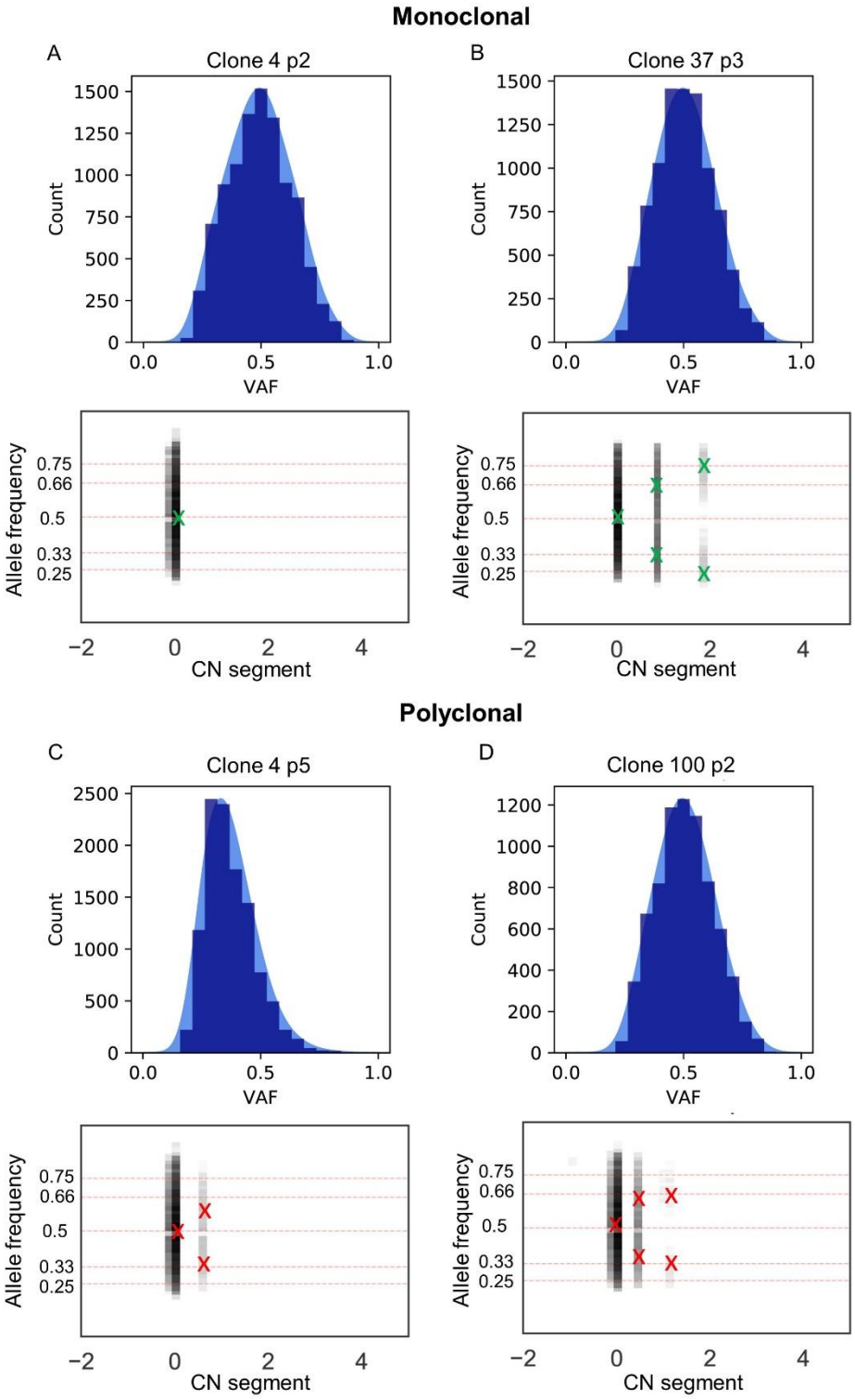


Figure 5.17: Clonality of samples determined by the VAF and CN method.
A) and B) examples of monoclonal populations with the green 'X' representing the peaks in each clone, indicative of ideal peaks associated with a monoclonal population. C) and D) examples of polyclonal populations with the red 'X' representing the peaks in each clone, indicative of distorted peaks associated with polyclonal populations. The top graphs represent the VAF method and the bottom graphs the CN method.

Category	Sample	Clonality status	Crisis duration
ALT-survived	Cl2 p2	Monoclonal	Short
	Cl2 p5	Monoclonal	
	Cl3 p2	Monoclonal	Short
	Cl3 p5	Monoclonal	
	Cl4 p2	Monoclonal	Short
	Cl4 p5	Polyclonal	
	Cl48 p5	Polyclonal	Long
	Cl131 p2	Monoclonal	Short
	Cl131 p7	Monoclonal	
	Cl147 p2	Monoclonal	Long
	Cl147 p5	Polyclonal	
ALT-died	Cl37 p3	Monoclonal	N/A
	Cl51 p2	Monoclonal	N/A
	Cl52 p4	Monoclonal	N/A
	Cl81p p2	Monoclonal	N/A
	Cl83 p2	Monoclonal	N/A
	Cl100 p2	Polyclonal	N/A
	Cl108 p2	Polyclonal	N/A
	Cl130 p5	Monoclonal	N/A
	Cl132 p3	Monoclonal	N/A
	Cl148 p5	Monoclonal	N/A
Telomerase +	Cl49 p2	Monoclonal	Short
	Cl49 p5	Monoclonal	
	Cl50 p5	Monoclonal	Long
	Cl53 p2	Polyclonal	Absent
	Cl53 p9	Monoclonal	
	Cl92 p4	Monoclonal	Long
	Cl103 p4	Monoclonal	Short
	Cl111 p6	Polyclonal	Short
	Cl112 p5	Monoclonal	Absent

Table 5.7: Summary of clonality status of each clone and correlation with crisis duration.

5.5 Discussion

5.5.1 Successful sequencing of telomere sequences

The data presented above show the successful sequencing of STELA products from the sub-telomere DNA to the telorette primer, which marks the end of the telomere. A good breadth of read lengths was obtained thus providing an interesting model for identifying unique events. However, a major limitation was the overall preference for the sequencing of short telomeres using the PacBio SMRT sequencing platform. A size selection step may be required for future use as recommended by PacBio, to maximise the number of long reads obtained. Although this method is associated with an error rate, similar to other sequencing platforms, it importantly allows the analysis of full telomere reads rather than attempting to assemble repetitive short contigs when using Illumina sequencing (Eid et al. 2009). Indeed, one current method used to sequence telomeres consists of Illumina sequencing of TTAGGG repeats to assess the error rate associated with false positive variant calling, followed by sequencing of telomeres of interest taking into account the error rate upon assembly of the sequenced telomeres (Lee et al. 2018). In addition, telomere variant repeat (TVR) PCR can be used to assemble the telomere according to the sequence of hexamers enabling allele separation between individuals (Baird et al. 1995). This technique was applied to ALT telomeres and compared to telomerase positive cells pre- and post-crisis to highlight differences in telomere sequence (Varley et al. 2002). ALT telomeres showed mutations at the start of the 16p/16q telomeres consistent with recombination events between telomeres. However, these techniques do not offer a complete accurate assembled long read which the PacBio SMRT sequencing does. This platform has been used in recent years to sequence highly repetitive GC-rich sequences, such as telomeres, with success but not in the context of ALT (Loomis et al. 2013; Lee et al. 2015b).

Therefore, to further characterise ALT elongation, more clones would need to be sequenced using SMRT sequencing for consistency and ideally, ALT telomeres would be sequenced before and after crisis to compare the effects of crisis on specific cells, as well as the upregulation of ALT which in turn may highlight a widespread ALT characteristic. One limitation is the ability to generate sufficient material as the library preparation for PacBio sequencing considerably reduces the material generated through the multiple steps involved, thus potentially limiting the pre-crisis sample as it can be difficult to obtain sufficient cells. Alternatively, the parental cell line could be used as a comparison as shown in this chapter. A large-scale analysis of multiple samples by the addition of tags may be

desirable to reduce cost and time. Indeed, tags, which consist of unique barcodes generated for single samples to differentiate these bioinformatically, can be included at the time of the PCR, which would require altering the primers used to amplify the DNA, but can also be attached to the dimers used during the library preparation of the amplicons prior to sequencing, post-PCR, although this requires considerable optimisation as it has been associated with demultiplexing issues upon processing the data (Buschmann et al. 2014). In addition, other novel sequencing platforms could be considered such as the MinION by Oxford technologies, as it similarly offers long sequencing reads, although this technique has encountered limited success with repetitive DNA and extensive optimisation would need to be carried out prior to use (Weirather et al. 2017).

5.5.2 ALT elongation is cell specific

Upon analysis of telomere variant repeats, a TTCGGG increase was seen in the HCT116 model whilst no increase in a specific variant repeat, but rather an overall increase in non-canonical variants, was observed in the HCA2 model but, interestingly, no specific increase in TTCGGG. These observations imply that ALT elongation is cell specific. In accordance with this hypothesis, Lee et al showed differences in variant repeats in two types of ALT positive cancer cell lines upon comparison to ALT negative controls: cells originating from melanoma and PanNETs (Lee et al. 2018). Both showed no statistical difference when comparing TTCGGG variant content. However, they did observe an overall increase in non-canonical variant repeats, similar to models studied here, with the exception of U2OS. This therefore shows that elongation of telomeres through the ALT mechanism is complex, and that variability is observed according to the cell type, rendering it difficult to observe similarities between cells, and further analysis will need to be undertaken before this information could be developed into diagnostic, prognostic or therapeutic tools (Lee et al. 2014).

The consistent increase of specific variants at multiple chromosome ends in the same cells indicates a common initial event that provides the template for the inserted sequences. The C-circles present in ALT cells may represent one potential candidate for the templates utilised for elongation of telomeres. This may arise through a rolling circle amplification, by which the single stranded DNA from the C-circle anneals to a strand of a telomere that requires elongation, followed by the synthesis of the sister strand (McEachern and Blackburn 1996; Tomaska et al. 2009). This would account for the sudden increase in length observed in the context of ALT (Teng et al. 2000). Therefore, sequencing these single stranded ECTRs could

address this question. For example, it will be informative in the HCT116 model that displays a stark increase in TTCGGG repeats if the C-circles presented with a high percentage of this specific variant. A hypothesis to explain the increase in a specific variant could be that the first C-circles generated, are then amplified exponentially to increase the amount of these ECTRs in each cell to contribute to the ALT phenotype. It is not entirely clear how these C-circles are generated but these could arise from a rolling circle amplification of a telomere that may contain a series of specific variant repeats, such as TTCGGG in the HCT116 cells, towards the start of the telomere, as we know that variant repeat variability is observed within the first 1 kb of the telomeric sequence (Baird et al. 1995; Baird et al. 2000). Another emerging hypothesis is that C-circles are by-products of BIR-mediated resolution of replication fork collapse therefore suggesting that replication stress contributes to the ALT phenotype (Zhang et al. 2019b). In addition, these C-circles appear early in the process of ALT initiation, as seen in chapters 3 and 4, further suggesting their importance early in the ALT mechanism.

Inter-chromosomal exchanges have been shown to contribute to telomere elongation by which short telomeres use long molecules as templates for HDR-mediated increase in length (Varley et al. 2002; Liu et al. 2018). Crisis is triggered by the accumulation of short telomeres and another hypothesis to explain the increase in a specific variant can be related to the telomeric sequence of the template telomere. Telomere dynamics vary at different chromosome ends, as shown in chapter 4 with long telomere observed at chromosome 8 (close to 6 kb) in HCT116 cells whilst critically short telomeres were measured at the XpYp telomeres (1 kb and below). Furthermore, telomere dynamics differ between cells, as seen with different STELA profiles in both chapters 3 and 4 upon comparison of different single cell clones. In addition, the specificity in insertion length has been hypothesised to be linked to a preference of exchanges between specific telomeres, perhaps due to the chromosomal location and compaction in the nucleus (Varley et al. 2002). Thus, the first telomere to trigger crisis in a cell uses a specific long telomere in proximity which may have an enrichment of a particular variant within the first kb of telomeric sequence. This in turn suggests that the shortest telomeres in each cell that trigger crisis may differ between clones and therefore the telomeric sequence will differ too, although, this would probably not account for the overall sudden increase in length or the widespread of a specific variant but may occur early and in combination with elongation using the C-circles as template.

5.5.3 ALT elongation is a result of multiple events

The data presented shows evidence that elongation mediated by the ALT mechanism results from multiple events as telomeres eroded to different lengths prior to undergoing elongation. Indeed, it appears that there is no definite length to which telomeres need to erode to prior to undergoing elongation, as seen with some telomeres being as short as 66 bps whilst others eroded to approximately 1 kb prior to elongation. It can however not be discounted that there may be a lower length threshold below which telomeres need to erode before they are subjected to elongation. Furthermore, no specific insert size was found in common between reads although ALT clones showed an increase in a specific variant repeat rendering it hard to estimate the length added after each elongation round.

In addition, it is not possible to predict whether telomeres are continually undergoing extension or if they erode and suddenly get elongated as it has previously been suggested (Teng et al. 2000). The heterogeneous telomere length distributions observed on HCA2 ALT clones and the U2OS ALT positive cancer cell line STELA blots could however be suggestive of a continual elongation irrespective of telomere length (Bryan et al. 1995; Henson et al. 2002). Additionally, in the HCT116 model, where homogeneous allelic distributions were observed and only one allele appeared to undergo elongation at the 17p chromosome on the STELA blots, it was anticipated that only one allele would show differences upon sequence analysis. This was not the case as both alleles at that chromosome end showed a considerable increase in TTCGGG suggesting that DNA is continually added to ALT telomeres, which further suggests that telomeres are continually elongated.

One hypothesis to explain the elongation of both alleles irrespective of telomere length could be related to the t-loop excision, described in section 1.3.5.4. The loss of ATRX results in increased replication stress and a less condensed telomeric chromatin due to the lack of incorporation of histone H3.3 (Episkopou et al. 2014). This in turn renders telomeres more exposed and prone to replication fork collapse and elicits the DDR. A long-standing hypothesis for the increase in HDR in ALT cells has been the excision of the t-loop formation, present at the end of chromosomes to protect the natural ends from the DDR, which generates blunt ended telomeres which become highly recombinogenic (Henson et al. 2002). This in turn signals to the DDR as the ends of chromosomes are recognised as DSBs and as ALT cells silence NHEJ-mediated repair, HDR is favoured for repair of the lesions and for extension thus resulting in different patterns in elongation (Koschmann et al. 2016). The t-

loop excision does not seem to occur at a specific length and may therefore be an indirect consequence of the loss of ATRX through its effects on telomere architecture.

5.5.4 Insertions and alterations in sequences

The insertion events observed in this dataset do not appear to be a characteristic of ALT. Indeed, prior to sequencing, one hypothesis to explain the chromosome specific elongation events observed in HCT116 cells was that a specific template DNA sequence was inserted into short telomeres to elongate these. If this piece of DNA was non-telomeric, this would have been observed in the data. Instead, the reads showed almost exclusively telomere repeats, with an increase in non-canonical variants in ALT cells, rendering it impossible to establish the length of insertion. In addition, the non-telomeric insertions that were observed appeared to be mostly G- and T-rich sequences, which were not mappable and suggestive of a sequencing artefact. The SMRT sequencing platform is associated with insertions and deletions of base pairs, although, the CCS reads used in this case usually reduce errors to obtain a 99% accuracy. Additionally, polymerase base substitution errors may also account for these differences, as SMRT sequencing uses real-time polymerase synthesis of the DNA being sequenced (Eid et al. 2009; Carneiro et al. 2012; Hestand et al. 2016). Finally, the occurrence of sub-telomere regions within the telomere would appear to be a result of rare recombination events during PCR or library preparation, as multiple chromosome ends were amplified in the same reactions, and both of these steps are involved with denaturation, synthesis and ligation of DNA (Potapov and Ong 2017).

The sub-telomeric DNA sequence was rarely altered and was always within 90% from the reference genome sequence, with minor insertions and deletions which can be associated with the SMRT sequencing platform (Eid et al. 2009; Carneiro et al. 2012). There was no evidence of erosion into the sub-telomere region combined with the generation of a new telomere which would result from NHEJ-mediated fusion events, which agrees with the hypothesis that ALT silences the NHEJ repair pathway (Koschmann et al. 2016). However, only reads that had a sub-telomere DNA that matched the chromosome ends amplified alongside a telorette sequence at the other end of the telomere were analysed. This means that reads that had an altered sub-telomere region (over 90% from expected sequence), were filtered out of the dataset, potentially dismissing real fusion events, although these are expected to be a result of crisis rather than the ALT mechanism, as fusions events decrease post-crisis in HCA2 cells and were absent in the U2OS cell line (chapter 4).

5.5.5 Loss of ATRX does not appear to alter the genome

The upregulation of ALT has been shown to increase instability in minisatellite MS32, located on chromosome 1, through inter-chromosomal exchanges which induce mutations (Jeyapalan et al. 2005). In addition, sarcomas are associated with a higher level of instability and chromothripsis (Cortes-Ciriano et al. 2018). Here, WGS of ALT positive HCT116 clones appeared to show more genome stability, than telomerase positive clones although, chromothripsis was observed in one sample (6.25% of ALT clones). Although only 16 ALT positive clones were sequenced, the rate of chromothripsis in this model is still higher than in the overall proportion of cancer cells (2-3%), although below the proportion seen in sarcomas (25%), thus suggesting the ALT mechanism may play a role in this. Undertaking WGS in the HCA2 fibroblast clones, which showed ALT upregulation exclusively for survival, would address this further as well as assess the impact the upregulation of ALT has on surviving cells and establish if chromothripsis is more common in ALT positive cells (Stephens et al. 2011).

Furthermore, although the overall genome was stable across all samples sequenced, a common gain was observed at chromosome 5 in 10 samples, some of which eventually upregulated telomerase for long-term survival. It is unclear why this was a common alteration although, HCT116 cells are cancerous and have therefore undergone a crisis phase and stabilised their genomes following escape from crisis (Green and Kaplan 2003). Single cell clones that share this common gain may have therefore originated from the same sub-population. Alternatively, the *hTERT* gene is located on the p arm of chromosome 5, close to the telomere, and amplification of the gene locus could be an adaptive mechanism these cells undergo to compensate for the downregulation of telomerase through DN-hTERT action to re-establish telomerase activity (Zhang et al. 2000). This may have also been the way these cells had upregulated telomerase initially and harboured this genomic change throughout the experiment. Once more, sequencing of HCA2 clones would be informative as these clones never upregulate telomerase for survival.

5.5.6 Conclusions

The loss if ATRX does not appear to have an impact on the overall genome compared with clones that upregulated telomerase for survival. In addition, differences in telomeric sequence were observed upon comparison of telomeres that had undergone ALT-like elongation and the parental controls. This was further shown to be cell specific as an increase in specific variant repeats was observed in each cell type, however, all ALT cells underwent an overall increase in non-canonical variant repeats. The elongation event is therefore common to all ALT cells but the sequence of said elongation is cell dependent and can perhaps be traced to the initial elongation event and dependent on the template used.

Chapter 6

General discussion

6.1 Summary

The aim of this thesis was to assess the impact the loss of ATRX has on the ability of cells to escape a telomere-driven crisis and how this influences genomic evolution, to ultimately gain a better understanding in the ALT mechanism.

Chapter 3 investigated how the loss of ATRX affected primary human fibroblast (HCA2) cells taken through a telomere-driven crisis. It was shown that loss of ATRX enabled cells to immortalise by upregulation of the ALT mechanism exclusively, and residual ATRX activity resulted in cell death. Telomeres were measured using STELA and showed a heterogeneous telomere length distribution prior to crisis which was increased post-crisis, consistent with an ALT upregulation. ALT positivity was confirmed by the presence of C-circles and the complete lack of telomerase activity. This showed that the combination of a telomere-driven crisis with the loss of ATRX is enough to activate and maintain ALT in primary human fibroblasts.

Chapter 4 examined the effect the loss of ATRX had on telomerase positive epithelial cancer cells (HCT116 cells). A telomere-driven crisis was initiated in these cells by downregulation of telomerase to induce telomere erosion and subsequently crisis. The results showed that the loss of ATRX hindered HCT116 cells' ability to escape crisis, whilst enabling the upregulation of the ALT mechanism for prolonged cellular division in a subset of clones (19%). In contrast to HCA2 cells, HCT116 cells showed homogeneous distributions prior to crisis, followed by a precise elongation event maintaining homogeneous telomere length distributions, suggestive of allele and chromosome specific elongation events. These clones also showed a rapid amplification of C-circles that was followed by death or telomerase activation; ultimately these clones were unable to maintain ALT activity. This did however show the ability of these cells to transiently switch to the ALT mechanism for telomere length maintenance for survival.

Chapter 5 aimed at analysing ALT telomere sequences alongside WGS to assess the impact the loss of ATRX and subsequently the upregulation of ALT had on the genome. Firstly, telomere sequences showed stark differences between ALT positive and normal or

telomerase positive controls. All ALT clones showed an increase in non-canonical variant repeats, with the HCT116 model showing a net increase in the TTCGGG variant across the three chromosome ends studied: 7p, XpYp and 17p. This provided evidence that telomeres from ALT cells differ greatly from telomeres being extended by telomerase and provided insight into the complexity of ALT elongation. It appeared that ALT elongation is cell type specific, and may even be clone specific, which adds to the variability of the ALT phenotype. WGS was carried out in HCT116 cells only, under multiple conditions: pre- and post-crisis in ALT-survived, ALT-died and telomerase positive clones. An overall increase in SVs was observed in telomerase positive cells and a higher CNV count in ALT positive cells, with the occurrence of chromothripsis in a unique ALT positive clone. Once more, ALT clones differed from telomerase positive clones further indicating that the genetic sequence is altered in the context of ALT and that may be of use to better understand the mechanism of ALT as well as develop a new diagnostic, prognostic and therapeutic tool to target such cancers.

6.2 Discussion

6.2.1 Loss of ATRX is sufficient to initiate the ALT mechanism

Firstly, in the HCA2 model, the complete loss of ATRX combined with telomere stress, through the abrogation of TP53 and Rb, showed replicative immortality in 100% of clones through the upregulation of the ALT mechanism, as seen with the lack of telomerase activity, a stark increase in C-circle intensity post-crisis and very heterogeneous telomere length distributions (Bryan et al. 1995; Henson et al. 2002). Mild residual ATRX activity was enough to inhibit survival thus confirming that ATRX is an inhibitor of the ALT mechanism (Napier et al. 2015). In addition, the loss alone of ATRX resulted in altered telomere length distributions as an overall increase in heterogeneity was observed across all clones that survived.

ATRX normally incorporates histone H3.3, alongside DAXX, to silence repetitive chromatin and therefore, the loss of ATRX affects telomere chromatin condensation and renders it more prone to transcription, as seen with the increase in TERRA in ALT cells (Goldberg et al. 2010; Episkopou et al. 2014). This therefore induces telomere dysfunction and ALT has been proposed as a mechanism to compensate for this dysfunction in the absence of ATRX (Li et al. 2019). However, due to an unsuccessful targeting of ATRX in HCA2 WT cells, it was difficult to assess whether the loss alone of ATRX was responsible for the increase in telomere length heterogeneity or, whether the closeness to crisis at the time of sampling was implicated in these observations. However, the downregulation of ATRX, as seen in some clones which had been targeted with the CRISPR, resulted in homogeneous telomere length distributions, resembling the controls, thus indicating the involvement of ATRX in telomere dynamics. Therefore, a disruption in the ATRX/DAXX/H3.3 complex may be one of the first steps in the upregulation of ALT.

ATRX has also been associated with resolution of secondary structures such as G-quadruplexes and R-loops that, if left unresolved, may pose a barrier to replication (Clynes et al. 2014; Nguyen et al. 2017). This further suggests that ATRX is lost early in the process of ALT as these secondary structures result in replication fork collapse, and resolution of these damage sites is carried out by HDR, and contributes to the ALT phenotype (Costes and Lambert 2012). In addition, this may result in highly recombinogenic telomeres, as the collapse of the replication fork exposes the ends of chromosomes to the repair machinery, and may result in the telomere length heterogeneity seen in these clones, and in other ALT positive cells, promoting HDR-mediated elongation of short telomeres. However, although replication stress is increased in ALT cells, it is thought to be somewhat limited as numerous

proteins compensate for the loss of ATRX (Min et al. 2017). Indeed, FANCM, for instance, is enriched at ALT telomeres as it resolves R-loops at those regions of the genome. A reduction of FANCM results in increased telomere dysfunction and a stark increase in C-circles, suggestive of an increased ALT positivity and the possibility that C-circles may be generated from resolution of replication stress (Pan et al. 2017; Pan et al. 2019; Zhang et al. 2019b). In addition, SMARCAL1 is abundant at telomeres that utilise the ALT mechanism for survival as it plays a major role in resolving replication stress at fragile sites such as telomeres. Loss of this protein resulted in increased DNA damage at telomeres and appears to be essential for survival of ALT cells (Cox et al. 2016). This implies that although replication stress is increased in ALT cells, a tight regulation is observed to ensure survival of cells.

To complement this initial experiment in fibroblasts, a telomere-driven crisis was initiated in telomerase positive epithelial cancer cells, by downregulating telomerase activity with a DN-hTERT which induces telomere erosion, to further assess how the loss of ATRX affects cells going through crisis. Unlike the HCA2 cells, ATRX loss hindered the cells' ability to escape crisis, as seen with a 22% survival rate, which was surprising as WT clones, with functional ATRX, all successfully escaped crisis by re-activation of telomerase (Jones et al. 2014). In addition, the ALT mechanism was upregulated transiently in 7 escaping clones, as telomerase was eventually re-activated for long-term survival and these cells do not appear to be able to maintain the ALT mechanism. In accordance with this, 21 clones were positive for C-circles but did not survive. Additionally, the loss of ATRX did not induce heterogeneous telomere length distributions, unlike the HCA2 cells, however, the rate of survival was greatly affected in these cells which may suggest an increased rate of telomere dysfunction due to the loss of ATRX and subsequently cell death, as the presence of C-circles was observed in some clones that died shortly after.

The interesting observation from this experiment was that despite a reduced rate of immortalisation, ALT was initiated in 19% of clones providing evidence that these cells can transiently upregulate ALT for prolonged proliferation. The reason for the differences observed between cell types is unclear, although, the developmental origin of these cells could account for the preference for the specific upregulation of a TMM. Primary human fibroblasts originating from MSCs show little telomerase activity during development, and tumours of mesenchymal origin have a higher propensity to utilise ALT. In contrast, epithelial cells show high levels of telomerase activity during development and a sustained activity in stem cells until differentiation into mature cells and epithelial-derived tumours

predominantly utilise telomerase (Wright et al. 1996; Zimmermann et al. 2003; Parsch et al. 2004).

Overall, the loss alone of ATRX in cells correlates with the appearance of ALT-associated hallmarks such as APBs and C-circles (Brosnan-Cashman et al. 2018). This could suggest that ALT is present at low levels in normal cells and the loss of its main repressor, ATRX, initiates the mechanism (Napier et al. 2015). This also suggests that these hallmarks appear early in the ALT mechanism but other events are required to fully establish this pathway for survival. This is further suggested by the fact that HCT116 cells, that had successfully initiated the generation of C-circles upon loss of ATRX and a telomere-driven crisis, ultimately died due to the inability to activate other events required for long-term ALT maintenance.

6.2.2 Insertion length consistency across models suggestive of a chromosome specific insertion

Intriguingly, telomere extension appeared to be chromosome specific in the HCT116 ALT clones. Indeed, consistent insertion lengths were measured across multiple clones whilst a significant difference was noticed between chromosome ends. Strikingly, the biggest insertion length was observed at the 17p chromosome ends, with an average of 5.43 kb. This led to the hypothesis that ALT elongation is chromosome specific and that a specific template is used at different chromosome ends. In the HCA2 model, no distinct allelic distributions were observed rendering it hard to assess the insertion length following elongation. Nonetheless, a bigger standard deviation was calculated after crisis at the 17p chromosome end, compared to the XpYp telomeres, indicating a wider range of telomere lengths thus suggesting a bigger insertion at 17p telomeres in HCA2 cells and further indicative of chromosome specific elongation.

Due to the inability to differentiate the long and short telomeres at the 17p chromosome end in the HCT116 model upon telomere sequencing, it was not possible to verify if a specific increase (5.43 kb) was observed. Furthermore, due to the increase in TTCGGG repeats and most telomeres consisting exclusively of that variant at the distal end, it was difficult to see a common insertion length and to establish if these variants were inserted at once or if it was a result of multiple extension events. The increase in this specific variant was observed at all alleles suggesting that ALT elongation affects all telomeres irrespective of telomere length. This would account for the heterogeneity observed in ALT positive cancer cells and questions

the hypothesis of erosion of telomeres followed by a sudden elongation event, although there may be a threshold length below which telomere need to erode to prior to elongation, as the 8q telomeres did not undergo elongation in the HCT116 clones as these were measured at around 6 kb (Teng et al. 2000).

If C-circles are used as templates for telomere elongation through rolling circle amplification, it is unclear how precise insertions occur at different chromosome ends as this appears to be a non-random event, nor why different chromosome ends undergo different elongation. The commonly used methods for measuring telomere length in the literature are terminal restriction fragment assays (TRFs), most recurring method for measuring ALT telomeres, and quantitative fluorescent in situ hybridisation, amongst others, which look at the overall population of telomeres (Bryan et al. 1995; Lai et al. 2018). Although these methods highlight the heterogeneity in telomere length in the context of ALT, STELA enables the analysis of single telomeres and offers a detailed analysis of telomere dynamics thus highlighting the differences in elongation at the chromosome ends studied (Baird et al. 2003). Therefore, perhaps further analysis of individual chromosome ends would provide more information regarding the chromosome specificity of ALT elongation and, attempting this in other ALT cells could address this aspect and would be a first step in attempting to understand the elongation event further. This does however show the complexity of ALT elongation and further characterisation of the mechanism used to elongate is required to attempt to understand the nature of the elongation sequence. Indeed, it has recently been shown that BIR is used for elongation in ALT cells but the template used for elongation remains speculative (Zhang et al. 2019a). Therefore, establishing whether C-circles through rolling-circle amplification and/or inter-chromosomal recombination occur is essential for a better understanding of ALT.

6.2.3 Fusions in ALT cells are associated with crisis

In HCA2 cells, fusions appeared when cells approached crisis with a maximal intensity recorded during deep crisis. The number of fusions then decreased as cells escaped crisis by ALT upregulation indicating that these fusions were a result of genomic rearrangements associated with crisis and not the ALT mechanism. This was confirmed by the complete lack of fusion events in the ALT positive U2OS cell line. In addition, ALT clones underwent more fusion events during crisis than control clones thus suggestive that these cells were more genetically unstable during that phase resulting in complex karyotypes (Guillou and Aurias

2010). Although this was not tested here, sarcomas, which show a higher proportion of ALT upregulation, exhibit complex karyotypes and cytogenetic changes consistent with higher rates of rearrangements during crisis and an increased rate of chromothripsis (Cortes-Ciriano et al. 2018). Although WGS was not undertaken in the HCA2 cells, results in HCT116 cells revealed a 6% rate of chromothripsis in ALT clones and this may be indicative of a correlation between ALT and chromothripsis, although this remains speculative and further WGS of HCA2 clones would address this further. This does signify that ALT positive cells are able to stabilise their genomes or limit rearrangements that occur from end-to-end fusions by silencing of NHEJ upon escape from crisis, despite an increase in telomere dysfunction in these cells (Koschmann et al. 2016; Li et al. 2019). Overall, these cells are transiently more prone to DNA damage, as seen with the increase in fusion events, but regain stability upon escape from crisis, as seen with the decrease in fusion events despite the presence of short telomeres seen on the STELA blots. In addition, HCT116 ALT positive clones showed a lack of SVs contrarily to telomerase positive cells further suggesting genome stability.

The analysis of sequenced telomeres did not highlight the presence of fusion events but rather showed an increase in non-canonical variant repeats, especially in the HCT116 model with the high proportion of TTCGGG repeats, notably at the distal end of the chromosomes studied. This in turn prevents Shelterin binding as TRF1 and TRF2 exclusively bind TTAGGG repeats, however, it enables other proteins to bind to these telomeres thus further altering the chromatin structure (Court et al. 2005; Conomos et al. 2012). In contrast, in the HCA2 and U2OS models, a higher increase in variants was seen but rarely at the distal end thus implying that Shelterin binding would be less affected. Nonetheless, this may render the chromatin available to cellular processes such as transcription and may signal to the DDR and favour HDR-mediated repair. This may also play a role in t-loop excision to generate T- and C-circles and induce a break to ensure elongation of the telomere through BIR (Henson et al. 2002).

6.2.4 Variability in ALT hallmarks

This thesis has highlighted the fact that variability in hallmarks is observed amongst different cell types that are ALT positive. Indeed, in the HCA2 cell line, heterogeneity in telomere length varied from one clone to the other, as seen with differences in standard deviations, thus implying a different effect on telomere dynamics between clones. In the HCT116 model, some cells displayed an elongation event whilst others only showed C-circle positivity.

Variation in intensity of C-circles was also observed and a decrease in intensity was seen upon co-existence of telomerase and ALT, suggestive that C-circle intensity directly correlates with levels of ALT activity (Henson et al. 2009; Pompili et al. 2017; Grandin et al. 2019). In addition, as previously mentioned, the cGAS-STING pathway has been recently shown to be dysfunctional in ALT positive cells thus allowing for the accumulation of ECTRs in the cytoplasm of cells (Chen et al. 2017). In both the HCA2 and HCT116 models, a downregulation of the STING protein was often observed in ALT clones, although not systematic, suggesting the involvement of other pathways to cope with the accumulation of C-circles. This once more provides evidence that ALT has a very heterogeneous phenotype and varies between clones. Moreover, the loss of TP53 facilitated the knock-out of ATRX and escape from crisis in fibroblasts resulting in ALT upregulation. The presence of WT TP53 in the HCT116 cells, did not appear to hinder the ability of cells to upregulate the ALT mechanism, however, these cells were unable to maintain ALT and a loss of TP53 may play a role in maintenance of ALT, although this remains to be fully tested (Bryan et al. 1997; Oppel et al. 2019). Finally, differences in telomere sequence were observed in ALT positive cancers. Indeed, although these all showed an increase in variant repeats, the increase in a particular variant was cell specific. This challenges the development of a robust assay for detecting ALT in tumour samples as well as limits the development of a new therapeutic tool.

6.2.5 Impact on research and clinical development

The increase in non-canonical variant repeats in the context of ALT suggests that the affinity of Shelterin binding, which exclusively binds TTAGGG repeats, may be reduced, thus potentially compromising the loop formation at the distal end (Court et al. 2005; Conomos et al. 2012). This implies that ALT telomere architecture is compromised resulting in an increase in telomere dysfunction and therefore prone to repair, notably through BIR which induces telomere elongation in the process (Li et al. 2019). This once more exhibits the heterogeneity encountered in ALT cells, as an increase in specific variants is observed in the dataset above and in the literature, further providing evidence that ALT hallmarks vary considerable between cell types and maybe between clonal populations of the same cell type (Lee et al. 2014; Lee et al. 2018). This further accentuates the need to understand the mechanism of elongation further to develop a robust testing for ALT positivity in cancer cells.

As mentioned above, what does appear to be consistent in ALT patients is the generation of C-circles. The gradual increase in C-circles in the HCA2 models and the variability in intensity

observed in the HCT116 cells shows evidence that levels of ALT activity can be determined by C-circle intensity (Henson et al. 2009; Grandin et al. 2019). Upon co-existence of telomerase and ALT, the levels of C-circle intensity dropped suggestive of a decrease in ALT activity. Therefore, this may offer a robust assay for detecting ALT positivity as well as stratify patients based on intensity. This in turn may be used as a prognostic marker as ALT positivity is associated with varying effects on survival rates. For instance, ALT positive glioblastomas show a better survival than ALT negative patients whilst most ALT positive STSs are associated with a poor outcome (Hakin-Smith et al. 2003; Liau et al. 2015a).

Upon downregulation of telomerase in epithelial cancer cells, a switch to the ALT mechanism was observed in 19% of clones. These cells showed the capacity to adapt to the downregulation of telomerase by activation of ALT which may compromise the efficiency of telomerase-targeted therapies currently being developed (Shay and Wright 2006; Hu et al. 2016). In contrast, fibroblast cells showed no telomerase upregulation in all clones studied providing evidence that cells that had initially upregulated telomerase may switch to ALT but not vice versa. In addition, some ALT positive cells were also positive for telomerase as cells re-activated telomerase for long-term survival indicating the possibility of both TMMs co-existing; although it is unclear if this would be maintained on the long-term as telomerase is thought to repress ALT (Perrem et al. 1999; Matsuo et al. 2009). However, it appears that the categorisation of cancers according to TMMs is not as simplistic as was once thought, as some samples show no signs of either known TMMs, indicating the possibility of a new TMM, whilst others show a transient co-existence which forms another barrier to targeting the process used to maintain telomeres for therapy (Costa et al. 2006; De Vitis et al. 2018).

What has shown promising results in ALT positive cells is the targeting of replication stress. Indeed, ALT cells undergo an increased rate of replication fork stalling and collapse due to the loss of ATRX which results in the accumulation of secondary structures (Clynes et al. 2014). This is limited by resolution of replication stress by HDR and the dissolution of these structures by other proteins such as FANCM (Costes and Lambert 2012; Ait Saada et al. 2018). However, an increase in replication stress, by inhibition of FANCM action, results in increased instability and cell death, thus implying the level of replication stress in ALT cells is finely balanced and contributes to the phenotype (Pan et al. 2017; Pan et al. 2019). Stabilisation of these G4 structures also induces cell death as the accumulation of unresolved secondary structures leads to a high rate of replication fork collapse and therefore death (Rizzo et al. 2009; Wang et al. 2019). Therefore, targeting the process of replication in ALT cells appears

to be the most promising avenue as of yet but a better understanding of the mechanism of elongation in ALT is required and may uncover another approach to targeting ALT in cancer.

6.3 Future directions

6.3.1 Knock out of ATRX

To further add to the dataset, the loss of ATRX should be tested in other primary human fibroblasts such as WI-38 or IMR-90, which are cell lines that are well-known and used in the literature. This experiment was also carried out in MRC-5 cells (unpublished data from our lab) showing similar features to the HCA2 cells presented above. It could therefore be predicted that other fibroblast cultures would behave in the same way. In addition, the MP in the HCA2 model showed early signs of crisis implying that the CRISPR transfection had affected the cells and a repetition of this experiment would be required. If cells escape, it would be interesting to assess TP53 status and, if unmutated, would show that its loss is not required but may facilitate the escape from crisis by upregulation of ALT. In addition, telomere lengths were highly heterogeneous prior to crisis in the HCA2 cells upon loss of ATRX. Undertaking the knockout of ATRX earlier in the lifespan of these cells would address whether the loss of ATRX is the cause for this or, if it is the loss of ATRX alongside the approach to crisis. If it is the loss of ATRX alone that causes these observations this would be an important indicator of the role this protein plays in telomere architecture. In addition, homogeneous distributions are observed in the HCT116 cells before and after crisis, despite undergoing ALT-like elongation, showing a stark difference to the HCA2 cells, and this could be indicative of a higher need for H3.3 incorporation in HCA2 cells and perhaps a compensatory mechanism in epithelial cells for the loss of ATRX.

Due to the potential involvement of TP53 in the successful long-term initiation of ALT, assessing the role of TP53 in the HCT116 model would be beneficial. Therefore, a repeat experiment with a double ATRX and TP53 knock out in these cells may further address the importance of TP53 in the ALT mechanism and establish whether cells would upregulate ALT more frequently in this context and more importantly, if they are capable of maintaining ALT on the long-term without switching to telomerase. This would also enlarge the dataset and further refine the percentage of ALT upregulation in epithelial cancer cells as well as further address the possible switch from telomerase to ALT when telomerase is downregulated, an important phenomenon in light of telomerase targeted therapies.

In addition, attempting to knock out ATRX in other cells of mesenchymal origin, such as osteoblasts, which show a high incidence of ALT upregulation and a high correlation of ALT with loss of ATRX, would further mimic the upregulation of ALT in cancer and further enlarge the dataset. Finally, studying cells from ATRX syndrome patients and taking these through

crisis would also be interesting to establish the incidence of replicative immortality in this context alongside the proportion of ALT positivity. Furthermore, analysis of telomere dynamics in these patients would address whether the loss alone of ATRX is responsible for the heterogeneity observed in the HCA2 cells.

6.3.2 Investigating HDR in ALT cells

As it has been mentioned extensively in the thesis, HDR plays an essential role in telomere elongation and more precisely the BIR pathway. Knocking out proteins involved in this mechanism, such as RAD51 and 52, in ALT positive cells, would provide more information on the mechanism of elongation. As previously mentioned, a better understanding of the mechanism of elongation is required for an overall better understanding of the ALT mechanism. Analysis of telomere dynamics using STELA in cells lacking major HDR proteins, such as RAD52 which appears essential, would provide insight in the elongation mechanism. In addition, to address whether C-circles are used for rapid elongation of telomeres, the prevention of the formation of these ECTRs, notably by the knock-out of BLM, would be insightful (Zhang et al. 2019a). Introducing tags within telomeres would also be informative as it may highlight common exchanges between pairs of chromosomes which may explain the specific insertion length observed in the ALT clones and establish if inter-chromosomal exchanges are responsible for telomere elongation and if this occurs between random chromosomes or if it is a tightly regulated phenomenon.

6.3.3 Whole genome sequencing

In this dataset only the HCT116 clones were sequenced for whole genome assembly to assess the effects of the loss of ATRX combined with telomere stress had on the overall genome. It was further used to compare an ALT upregulation with a telomerase upregulation. To add to this, WGS before and after crisis of HCA2 ALT-positive clones would be useful to establish if the lack of SVs observed in the HCT116 model is related to ALT or if it is specific to the HCT116 model. It would also highlight potential chromothriptic samples and establish if there is a higher incidence of this phenomenon in ALT positive clones compared to ALT negative clones and whether the higher incidence of chromothripsis in sarcomas occurs due to ALT upregulation. Furthermore, due to some ALT positive cancers retaining WT ATRX activity, the

identification of other underlying mutations in the context of ALT would be interesting and necessary for uncovering the involvement of other pathways that appear to be essential in ALT maintenance especially.

6.3.4 Telomere sequencing

Further sequencing of telomeres of ALT clones pre- and post-crisis would further confirm whether the events seen here, i.e. the increase in TTCGGG variant repeat in the HCT116 ALT clones, are general across all clones of a same model, or if it is clone specific. In addition, comparing the same clone before and after crisis would show a direct change in telomere repeat content as a result of crisis and upregulation of ALT. To further address whether the increase in a particular variant repeat is due to using C-circles as templates, sequencing of amplified C-circles for each respective clone would address this question. It would also address the hypothesis that the initial elongation events result from using C-circles as templates.

6.4 Conclusions

6.4.1 Project conclusions

The data presented in this thesis showed:

- Complete loss of the ATRX protein results in replicative immortality by upregulation and maintenance of ALT in primary fibroblasts when TP53 is dysfunctional as seen by the generation of C-circles, lack of telomerase and heterogeneous telomere lengths
- Loss of ATRX is enough to initiate but not maintain the ALT mechanism in a proportion of telomerase positive epithelial cells upon downregulation of telomerase activity as seen with the generation of C-circles, lack of telomerase and telomere elongation
- Re-establishment of telomerase activity in epithelial cells for long-term survival shows a preference to activate telomerase to achieve replicative immortality but the transient ability adapt to telomerase inhibition by using ALT if required. This is an important consideration for the development of telomerase targeted therapy, but also suggests other underlying events involved in maintenance
- Telomere elongation through ALT appears to be chromosome specific and affects all alleles irrespective of length, although telomeres may need to reach a certain threshold prior to undergoing elongation
- ALT elongation is a result of multiple events as telomeres showed different rates of erosion prior to undergoing an ALT-like elongation
- Telomere sequence analysis of ALT clones showed an overall increase in non-canonical variant repeats with a specific variant increase dependent on cell type and potentially on the first elongation event
- WGS showed a more stable genome in ALT positive clones versus telomerase positive clones although one ALT clone showed signs of chromothripsis

6.4.2 General conclusion

ATRX is a strong repressor of ALT activity and is often mutated in ALT positive cancers and its mutation therefore appears to be important but not essential as some tumours retain ATRX activity. The loss of ATRX is associated with heterogeneous telomere length distributions involving ATRX in telomere function and architecture. Indeed, ATRX functions at telomeres by incorporation of histone H3.3 which condenses the chromatin and silences it preventing

transcription of telomeres into TERRA. The loss of ATRX is therefore associated with decompressed chromatin rendering it more available for transcription and prone to R-loop formations. This in turn results in increased replication stress and HDR-mediated repair resulting in the increased rate of recombination associated with the ALT phenotype. C-circles and APBs appear early in the process of ALT, following ATRX loss and induction of crisis, and this is suggestive of their involvement in telomere elongation. This appears to be sufficient to initiate the ALT mechanism, however, other events are required for maintenance. Furthermore, the involvement of TP53 may facilitate the upregulation of the ALT mechanism but may not be essential. A proposed mechanism for ALT elongation is presented in figure 6.1 and combines the data presented in this thesis with the literature.

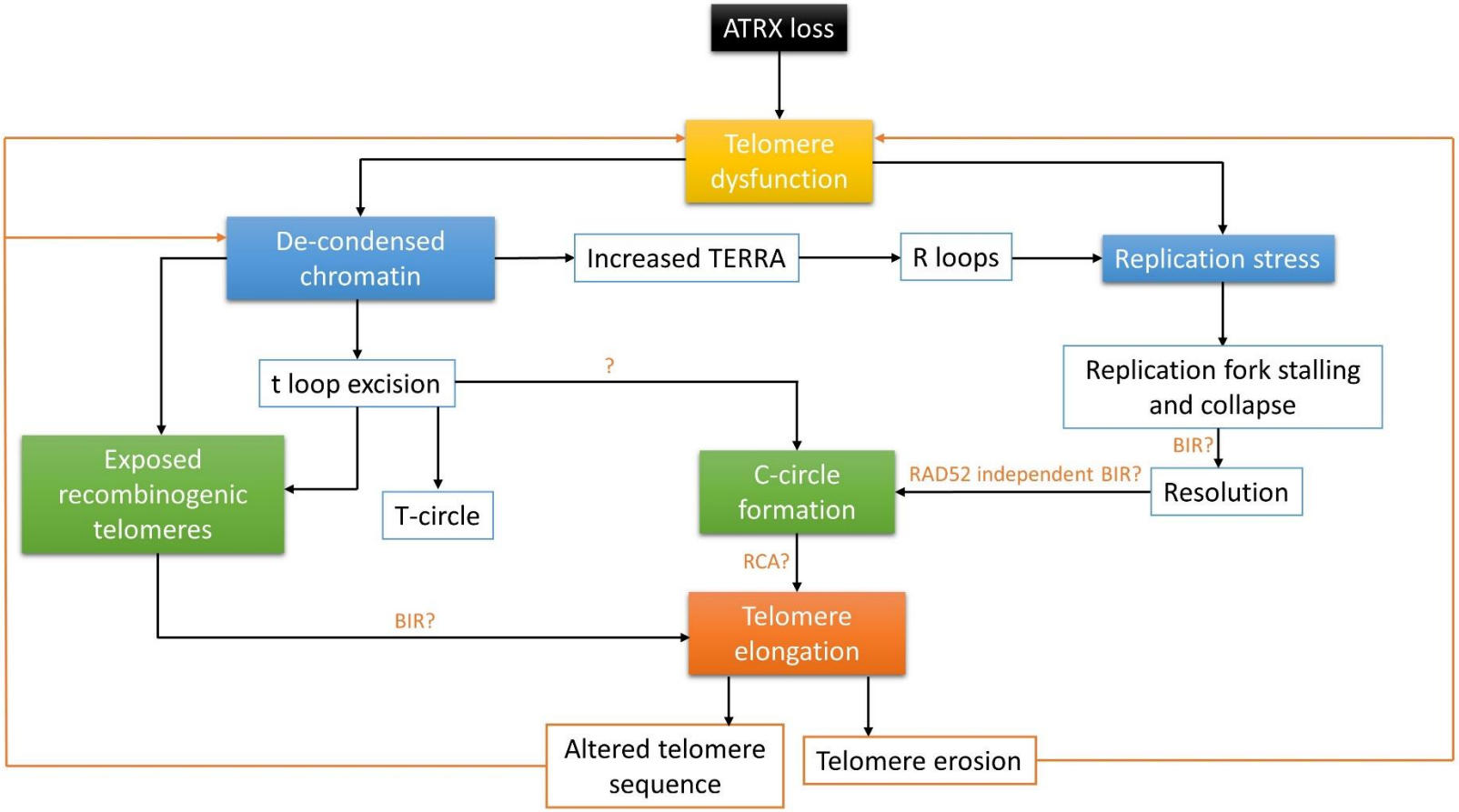
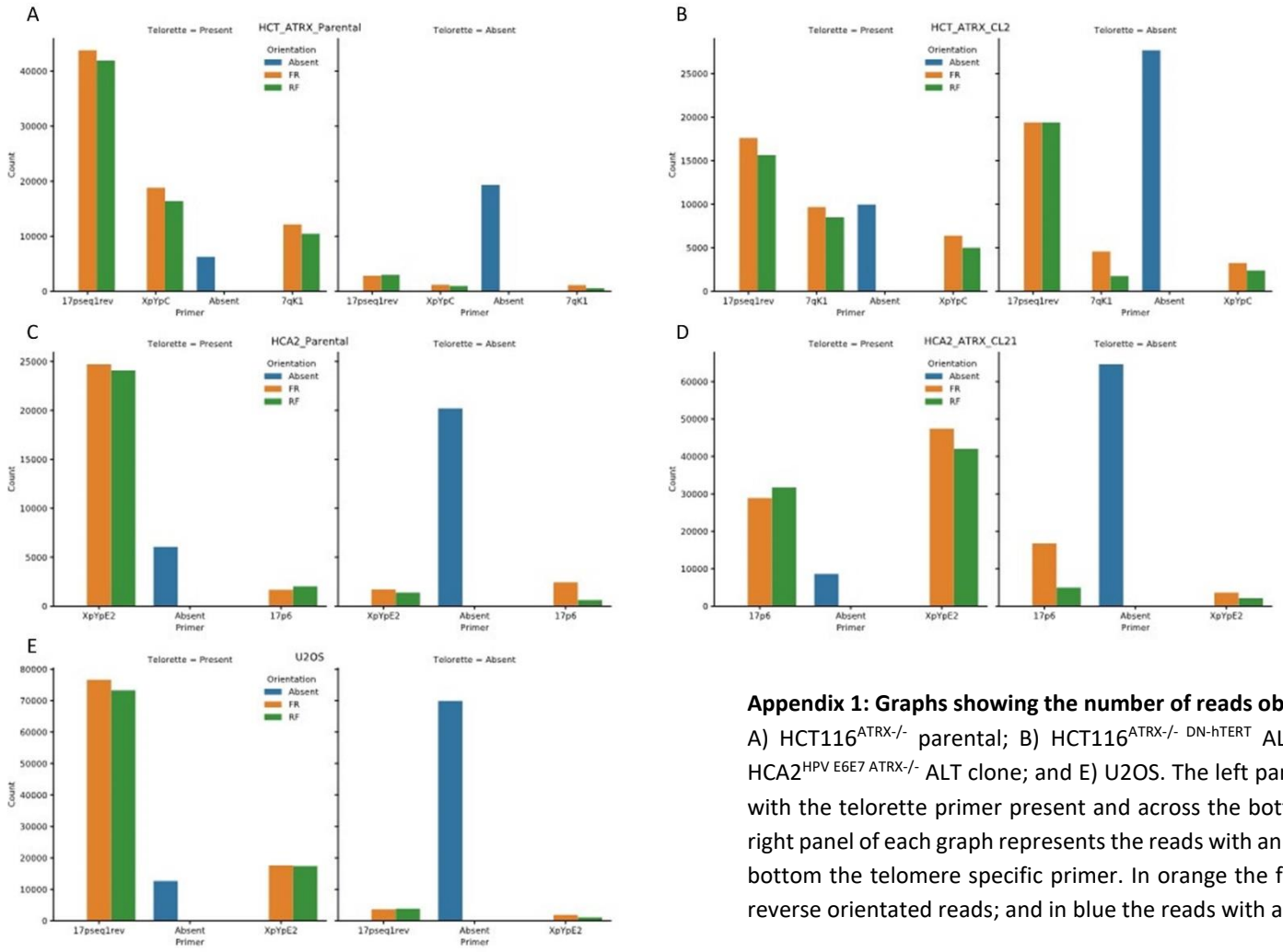


Figure 6.1: Proposed hypothesis for telomere elongation through ALT upregulation upon loss of ATRX.
The loss of ATRX leads to telomere dysfunction and induces a modified telomere architecture leading to exposed telomeres and t-loop excision which in turn generates T- and maybe C-circles. Loss of ATRX also induces replicative stress due to an accumulation of secondary structures and the resolution of these also results in the generation of C-circles. C-circles may be used as templates for telomere elongation through RCA as well as exposed telomeres which use other telomeres as templates. Telomere architecture is then altered due to an increase in non-canonical variant repeats and leads to further telomere dysfunction. Telomeres also erode and become dysfunctional.

APPENDICES



Appendix 1: Graphs showing the number of reads obtained through PacBio sequencing. A) HCT116$^{ATRX-/-}$ parental; B) HCT116$^{ATRX-/-\ DN\text{-}hTERT}$ ALT clone; C) HCA2$^{HPV\ E6E7}$ parental; D) HCA2$^{HPV\ E6E7\ ATRX-/-}$ ALT clone; and E) U2OS. The left panel of each graph represents the reads with the telorette primer present and across the bottom the telomere specific primer. The right panel of each graph represents the reads with an absent telorette primer and across the bottom the telomere specific primer. In orange the forward orientated reads; in green the reverse orientated reads; and in blue the reads with an absent telomere specific primer.

Number101113

TGCTCCGTGCATCTGGCATCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTGACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACCCTAACC
CTAACCAGGGACAGCCCGTTGCATGAGCCC
CTGGACTCTAACCCAGTTCCCTTCTGGAATTTAGGGGCCCTGGGACAGCCCTGTACATGAGCTCCTGGTCTGTAACACAGTTCCCCTGTGGGGATTTAGGGACTTGGGCCTTCTGTCTTTGGGATCTACTCTCTATGGGCCACACAGATATGTCTTCCAACTTCCCT
ACACAGGGGGGACTTCAAAGAGTGCCTTGAGCTGATCTGGTGATTGCTTTTTGTACTGTTATTTATCTTATTCTTTTCATTGTGAGGTACTGATGCAAACACTTTGTACGAAAAGGTCTTTCTCATCTCGGGAGTCCCCGTCTATTTGTCCCGGTCCCTG

0000000000000000111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111
11111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111111000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000

Number33034

CAGGGACCGGGACAAATAGACGGGGACTCCCGAGATGAGAAAGACCTTTTCGTACAAAGTGTTTGCATCAGTACCTCACAATGAAAAGAATAAGATAAATAACAGTACAAAAAAGCAATCACCAGATCAGCTCAAGGCACTCTTTGAAGTCCCCCCTGTGTAGG
GAAGTTGGAAGACATATCTGTGTGGCCCATAGAGAGTAGATCCCAAAGACAGAAGGCCCAAGTCCCTAAATCCCCACAGGGGAACTGTGTTACAGACCAGGAGCTCATGTACAGGGCTGTCCCAGGGCCCCTAAATTCCAGAAGGGAACTGGGTTAGAGTCCA
GGGGCTCATGCAACGGGCTGTCCCTGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTGGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGGTTAGGGTTAGGGGGTTAGG
GGTTAGGGTTAGGGTTAGGGTTAGGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGGTGAGGGGTTAGGGTTAGGGTTAGGGGTTAGGGTTAGGGGTTAGGGTTAGGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTT
AGGGTTAGGGT
TAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGATGCCAGATGCACGGAGCA

000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
0000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000002222222
2222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
2222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
22222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222000000000000000

Number54925

CAGGACCGGACAATAGACGGGGACTCCCGAGATGAGAAAGACCTTTTCGTACAAAGTGTTTGCATCAGTACCTCACAATGAAAGAATAAGATAAATAACAGTACAAAAAAGCAATCACCAGATCAGCTCAAGGCACTCTTTGAAGTCCCCCCTGTGTAGGGAAG
TTGGAAGACATATCTGTGTGGCCCATAGAGAGTAGATCCCAAAGACAGAAGGCCCAAGTCCCTAAATCCCCACAGGGGAACTGTGTTACAGACCAGGAGCTCATGTACAGGGCTGTCCCAGGGCCCCTAAATTCCAGAAGGGAACTGGGTTAGAGTCCAGGGG
CTCATGCAACGGGCTGTCCCTGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGGAGGGGGAGGGGGAGGGTGAGGGGGAGGGGGAGGGGGAGGGGGAGGGG
GAGGGGGAGGGGGAGGGGGAGGGGGAGGGCGAGGGGGCAGGGGGAGGGGGAGGGGAGGGGGAGGGGGAGGGGGAGGGGGAGGGGAGGGGGAGGGGGAGGGTGGAGGGTGAGGGGGAGGGGGAGGGGAGGGGGAGGGGGAGGGTGA
GGGGGTAGGGGGAGGGGGAGGGGTAGGGGGTAGGGGTAGGGGTAGGGGTAGGGGTAGGGTGAGGGGTGGTGAGGGGGAGGGGGAGGGGGAGGGTTAGCGGTTAGGGGGAGGGGTAGGGGTAGGGTGAGGGGAGGGTGTAGGGTTAGGGGTAGGG
TTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGGTAGGGTTAGGGTTAG
GGTTAGGGTAGGTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGATGCCCAGATGCACGGAGCA

000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000022222222222
2222222222222222222222222222222222222222222222222222222222222222222222222222222000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000002222222222222222222222222222222222222222222222000000000000000000000
00000022222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222
22222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222222000000000000000000

Appendix 2: Examples of reads generated using the hidden Markov model with the original sequence above.
In green the telomere specific primer sequence; in dark blue the sub-telomere sequence; in light blue the telomere repeats, forward (1) and reverse orientation (2); in orange end insertions; in red interstitial insertions; and in purple the telorette primer.

HCT116$^{ATRX-/-}$

XpYp telomere

TGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGT

17p telomere Allele 1

AAGGGTTAGGGTTAGGGTTGGGGTTGGGGTTGGGGTTGGGGTTAGGCTTAGGGCTTAGGGCTTAGGGCTAGGGCTAGGGCTAGGGCTAGAGTTAGGGTTG

17p telomere Allele 2

GAAGGGTTAGGGTTAGGGTTAGGGTTGGGGTTGGGGTTGGGGTTAGGGTTAGGGTTGGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAG

7q telomere Allele 1

CAGTGTTAGGGTTAGGGGTTAGGGGTTAGGGTCAGGGTCAGGGTCAGGGTCAGGGTCAGGGTTAGGGTAGGGTTCGGGTTCGGGTTCGGGTTCGGGTTCG

7q telomere Allele 2

CAGTGTTAGGGTTAGGGGTTAGGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGGTTAGGGGTTGGGGTTGGGGTTAGGGGTTAGGGGTTGGG

HCA2$^{HPV\ E6E7}$

XpYp telomere Allele 1

GAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTGGGGTTGGGGTTGGGGTTGGGGTTAGG

XpYp telomere Allele 2

GAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTTAGGGTCAGGGTTAGGGGTTAGGGTCAGGGTTAGGGGTTAGGGTCAGGGTCAGGGTCAGGGTC

17p telomere Allele 1

GAAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTGGGGTTGGGGTTGGGGTTGGGGTTGGGGTTGGGGTTGGGGTTGGGGTTGGGGTTAG

17p telomere Allele 2

AAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTGGGGTTGGGGTTGGGGTTAGGGTTAGGGTTGGGGTTAGGGTTGGGGTTAGGGTTAGGGTTAGGGTTAGG

U2OS

XpYp telomere

GAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGG

17p telomere Allele 1

AAGGGTTAGGGTTAGGGTTGGGGTTGGGTTGGGGTTGGGGTTAGGCTTAGGGCTTAGGGCTTAGGGCTAGGGCTAGGGCTAGGGCTAGAGTTAGGGTTGG

17p telomere Allele 2

GAAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTGGGGTTGGGGTCGGGGTTAGGGTTAGGGTCGGGGTCAGGGTCAGGGTCAGGGTCAGGGTCAGGGTCAG

Appendix 3: Examples of the highest scoring reads for each allele at each chromosome end studied.

The parental samples to represent the HCT116$^{ATRX-/-}$ and HCA2$^{HPV\ E6E7}$ cell lines and the U2OS cell line. Legend as follows: TTAGGG; TCAGGG; TTCGGG; GTAGGG; TGAGGG; TTGGGG; TAAGGG; CTAGGG; TTTGGG; TGCGGG; AGAGGG.

HCT116$^{ATRX-/-}$

XpYp

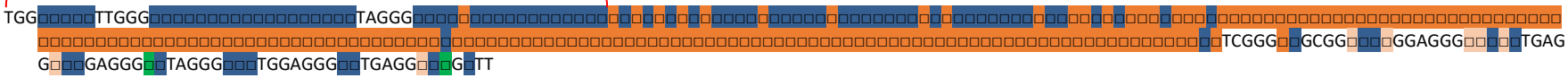
- Parental (Reference)

- ALT clone

(unaltered)

225 bps

253 bps



69 bps

7q allele 1

- Parental (reference)

- ALT clone

(unaltered)

113 bps

434 bps

711 bps

7q allele 2

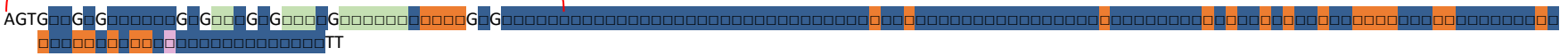
- Parental (reference)

- ALT clone

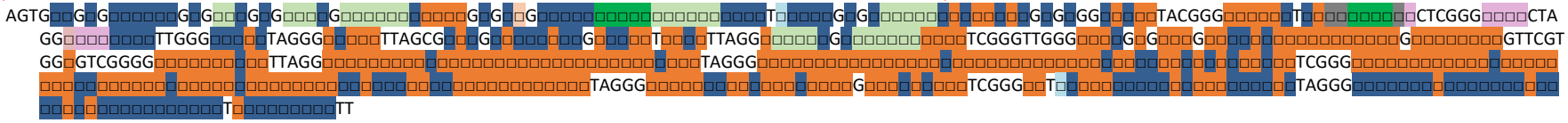
(unaltered)

66 bps

223 bps



347 bps



HCA2$^{HPV\ E6E7}$

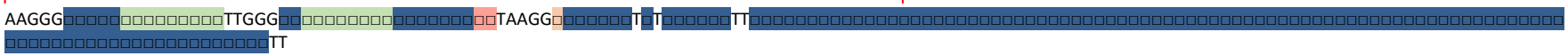
17p allele 1

- Parental (reference)

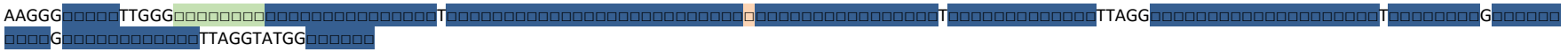
- ALT clone

(unaltered)

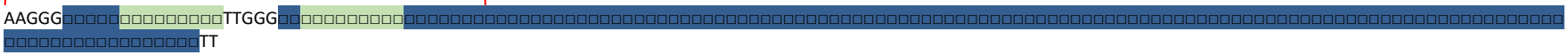
379 bps

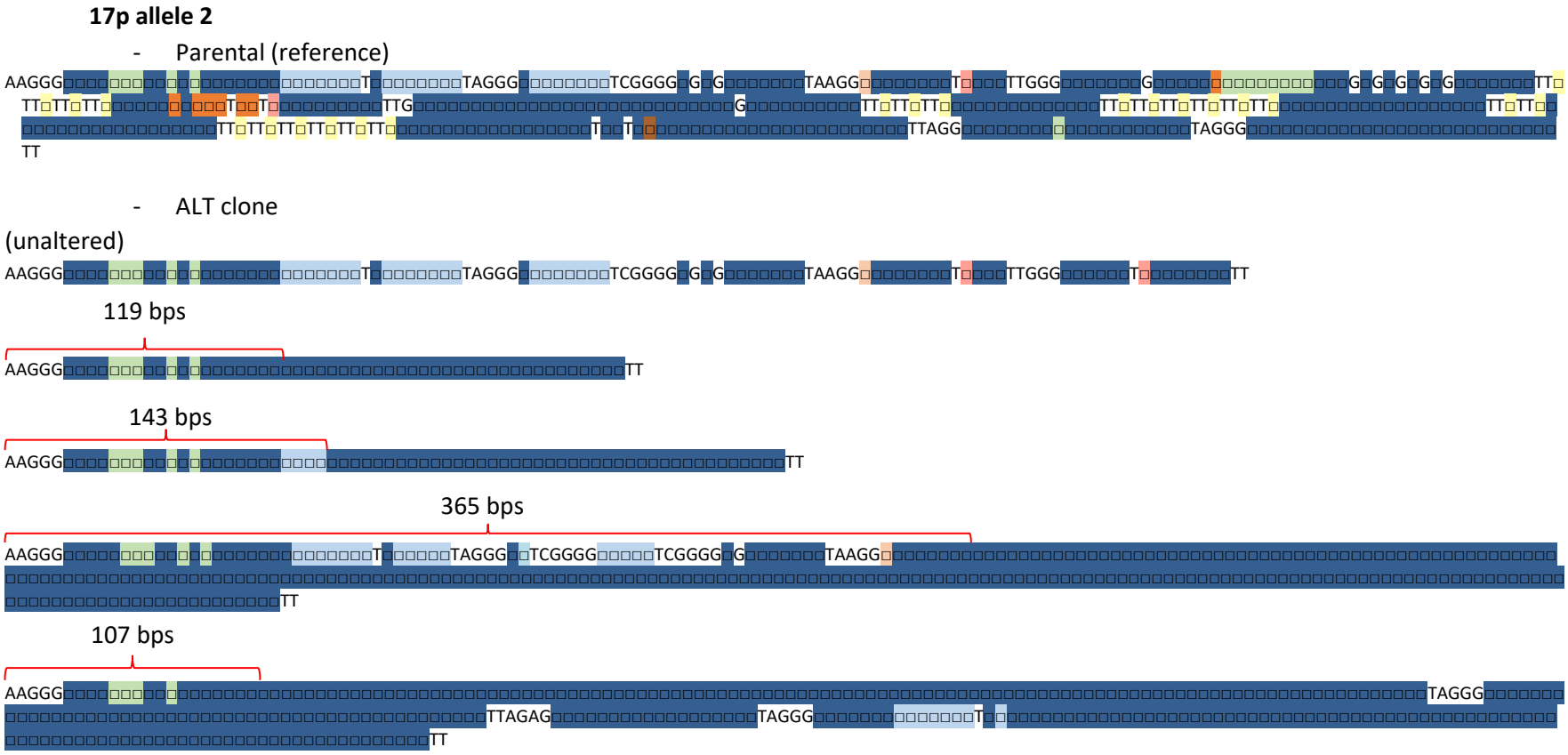


100 bps



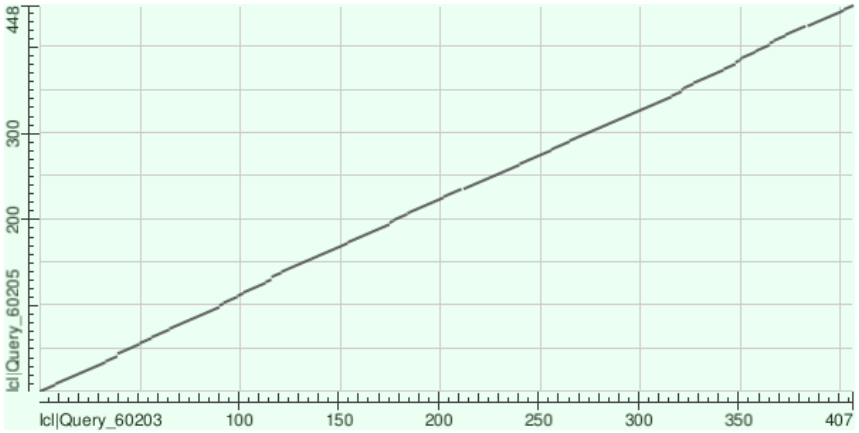
202 bps



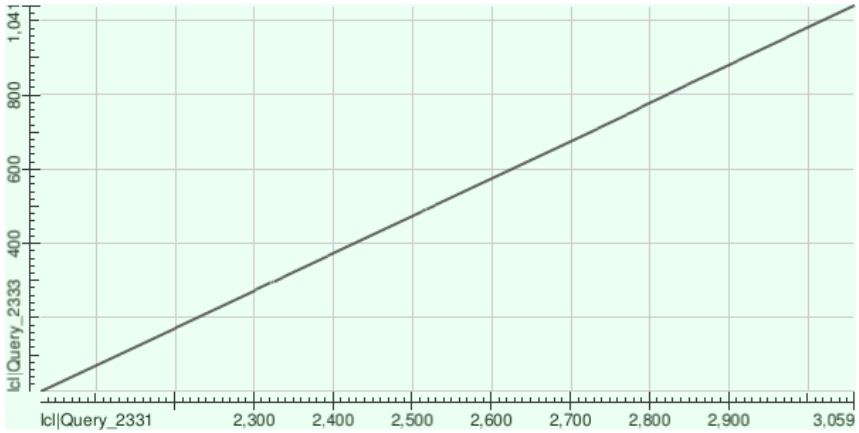


Appendix 4: Examples of multiple elongation events in all samples at all chromosome ends studied.
The HCT116$^{ATRX-/-}$ parental and the HCT116$^{ATRX-/-\ DN\text{-}hTERT}$ ALT clone at the XpYp and 7q chromosome ends; and HCA2$^{HPV\ E6E7}$ parental and the HCA2$^{HPV\ E6E7\ ATRX-/-}$ ALT clone at the 17p chromosome end. In red the estimated length to which telomeres eroded to before undergoing extension upon comparison to the respective parental sequences. Legend as follows: TTAGGG; TCAGGG; TTCGGG; GTAGGG; TGAGGG; TTGGGG; TAAGGG; CTAGGG; TTTGGG; TGCGGG; AGAGGG. □ represents a 6 nucleotide variant repeat and non-highlighted sequences represent single nucleotides.

A



B



C

D

Appendix 5: Examples of altered sub-telomeres (more than a 10% change from expected length) of reads for which the primer and the sub-telomere sequence matched. A) Dot plot result comparing the expected sub-telomere sequence (X-axis) against the altered sub-telomere sequence (y-axis; 41 bps insertion) for the HCA2$^{HPV\ E6E7\ ATRX-/-}$ ALT clone at the XpYp allele 1. B) Dot plot result comparing the expected sub-telomere sequence (X-axis) against the altered sub-telomere sequence (y-axis; 2,017 bps deletion) for the HCA2$^{HPV\ E6E7\ ATRX-/-}$ ALT clone at the 17p allele 2. C) and D) the telomere to the sub-telomere regions analysed matching A) and B) respectively with the expected telomere at the top and the telomere taken from the altered sub-telomere at the bottom. Dot plots generated on BLAST. Legend as follows: TTAGGG; TCAGGG; TTCGGG; GTAGGG; TGAGGG; TTGGGG; TAAGGG; CTAGGG; TTTGGG; TGCGGG; AGAGGG. □ represents a 6 nucleotide variant repeat and non-highlighted sequences represent single nucleotides.

REFERENCES

Abedalthagafi, M. et al. 2013. The alternative lengthening of telomere phenotype is significantly associated with loss of ATRX expression in high-grade pediatric and adult astrocytomas: a multi-institutional study of 214 astrocytomas. *Mod Pathol* 26(11), pp. 1425-1432. doi: 10.1038/modpathol.2013.90

Acosta, J. C. et al. 2013. A complex secretory program orchestrated by the inflammasome controls paracrine senescence. *Nat Cell Biol* 15(8), pp. 978-990. doi: 10.1038/ncb2784

Ait Saada, A. et al. 2018. Preserving replication fork integrity and competence via the homologous recombination pathway. *DNA Repair (Amst)* 71, pp. 135-147. doi: 10.1016/j.dnarep.2018.08.017

Aksenova, A. Y. et al. 2013. Genome rearrangements caused by interstitial telomeric sequences in yeast. *Proc Natl Acad Sci U S A* 110(49), pp. 19866-19871. doi: 10.1073/pnas.1319313110

Allfrey, V. G. et al. 1964. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proc Natl Acad Sci U S A* 51, pp. 786-794. doi: 10.1073/pnas.51.5.786

Allshire, R. C. et al. 1989. Human telomeres contain at least three types of G-rich repeat distributed non-randomly. *Nucleic Acids Res* 17(12), pp. 4611-4627. doi: 10.1093/nar/17.12.4611

Ardui, S. et al. 2018. Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Res* 46(5), pp. 2159-2168. doi: 10.1093/nar/gky066

Arora, R. et al. 2014. RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nat Commun* 5, p. 5220. doi: 10.1038/ncomms6220

Asamitsu, S. et al. 2019. Recent Progress of Targeted G-Quadruplex-Preferred Ligands Toward Cancer Therapy. *Molecules* 24(3), doi: 10.3390/molecules24030429

Atkinson, S. P. et al. 2005. Lack of telomerase gene expression in alternative lengthening of telomere cells is associated with chromatin remodeling of the hTR and hTERT gene promoters. *Cancer Res* 65(17), pp. 7585-7590. doi: 10.1158/0008-5472.can-05-1715

Avilion, A. A. et al. 1996. Human telomerase RNA and telomerase activity in immortal cell lines and tumor tissues. *Cancer Res* 56(3), pp. 645-650.

Aviv, A. et al. 2011. Impartial comparative analysis of measurement of leukocyte telomere length/DNA content by Southern blots and qPCR. *Nucleic Acids Res* 39(20), p. e134. doi: 10.1093/nar/gkr634

Azzalin, C. M. et al. 2007. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* 318(5851), pp. 798-801. doi: 10.1126/science.1147182

Baird, D. M. 2018. Telomeres and genomic evolution. *Philos Trans R Soc Lond B Biol Sci* 373(1741), doi: 10.1098/rstb.2016.0437

Baird, D. M. et al. 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* 66(1), pp. 235-250. doi: 10.1086/302721

Baird, D. M. et al. 2004. Normal telomere erosion rates at the single cell level in Werner syndrome fibroblast cells. *Hum Mol Genet* 13(14), pp. 1515-1524. doi: 10.1093/hmg/ddh159

Baird, D. M. et al. 1995. Mechanisms underlying telomere repeat turnover, revealed by hypervariable variant repeat distribution patterns in the human Xp/Yp telomere. *Embo j* 14(21), pp. 5433-5443.

Baird, D. M. et al. 2003. Extensive allelic variation and ultrashort telomeres in senescent human cells. *Nature Genetics* 33(2), pp. 203-207. doi: 10.1038/ng1084

Balk, B. et al. 2013. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nat Struct Mol Biol* 20(10), pp. 1199-1205. doi: 10.1038/nsmb.2662

Bannister, A. J. and Kouzarides, T. 2011. Regulation of chromatin by histone modifications. *Cell Res* 21(3), pp. 381-395. doi: 10.1038/cr.2011.22

Barefield, C. and Karlseder, J. 2012. The BLM helicase contributes to telomere maintenance through processing of late-replicating intermediate structures. *Nucleic Acids Res* 40(15), pp. 7358-7367. doi: 10.1093/nar/gks407

Baumann, P. and Cech, T. R. 2001. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* 292(5519), pp. 1171-1175. doi: 10.1126/science.1060036

Bayani, J. et al. 2007. Genomic mechanisms and measurement of structural and numerical instability in cancer cells. *Semin Cancer Biol* 17(1), pp. 5-18. doi: 10.1016/j.semcancer.2006.10.006

Benetos, A. et al. 2018. Short Leukocyte Telomere Length Precedes Clinical Expression of Atherosclerosis: The Blood-and-Muscle Model. *Circ Res* 122(4), pp. 616-623. doi: 10.1161/circresaha.117.311751

Benetos, A. et al. 2019. Telomere length tracking in children and their parents: implications for adult onset diseases. *Faseb j* 33(12), pp. 14248-14253. doi: 10.1096/fj.201901275R

Berardinelli, F. et al. 2010. Transient activation of the ALT pathway in human primary fibroblasts exposed to high-LET radiation. *Radiat Res* 174(5), pp. 539-549. doi: 10.1667/rr2127.1

Bergmann, O. et al. 2015. Dynamics of Cell Generation and Turnover in the Human Heart. *Cell* 161(7), pp. 1566-1575. doi: 10.1016/j.cell.2015.05.026

Bertrand, P. et al. 1997. Increase of spontaneous intrachromosomal homologous recombination in mammalian cells expressing a mutant p53 protein. *Oncogene* 14(9), pp. 1117-1122. doi: 10.1038/sj.onc.1200931

Bertulat, B. et al. 2012. MeCP2 dependent heterochromatin reorganization during neural differentiation of a novel Mecp2-deficient embryonic stem cell reporter line. *PLoS One* 7(10), p. e47848. doi: 10.1371/journal.pone.0047848

Berube, N. G. 2011. ATRX in chromatin assembly and genome architecture during development and disease. *Biochem Cell Biol* 89(5), pp. 435-444. doi: 10.1139/o11-038

Berube, N. G. et al. 2002. Neurodevelopmental defects resulting from ATRX overexpression in transgenic mice. *Hum Mol Genet* 11(3), pp. 253-261. doi: 10.1093/hmg/11.3.253

Bilaud, T. et al. 1997. Telomeric localization of TRF2, a novel human telobox protein. *Nat Genet* 17(2), pp. 236-239. doi: 10.1038/ng1097-236

Blackburn, E. H. and Gall, J. G. 1978. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. *J Mol Biol* 120(1), pp. 33-53. doi: 10.1016/0022-2836(78)90294-2

Bohr, V. A. 2002. Human premature aging syndromes and genomic instability. *Mech Ageing Dev* 123(8), pp. 987-993. doi: 10.1016/s0047-6374(02)00039-8

Bolzan, A. D. 2017. Interstitial telomeric sequences in vertebrate chromosomes: Origin, function, instability and evolution. *Mutat Res* 773, pp. 51-65. doi: 10.1016/j.mrrev.2017.04.002

Bower, K. et al. 2012. Loss of wild-type ATRX expression in somatic cell hybrids segregates with activation of Alternative Lengthening of Telomeres. *PLoS One* 7(11), p. e50062. doi: 10.1371/journal.pone.0050062

Britt-Compton, B. et al. 2006. Structural stability and chromosome-specific telomere length is governed by cis-acting determinants in humans. *Human Molecular Genetics* 15(5), pp. 725-733. doi: 10.1093/hmg/ddi486

Broccoli, D. et al. 1997. Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nat Genet* 17(2), pp. 231-235. doi: 10.1038/ng1097-231

Broccoli, D. et al. 1995. Telomerase activity in normal and malignant hematopoietic cells. *Proc Natl Acad Sci U S A* 92(20), pp. 9082-9086. doi: 10.1073/pnas.92.20.9082

Brosnan-Cashman, J. A. et al. 2018. ATRX loss induces multiple hallmarks of the alternative lengthening of telomeres (ALT) phenotype in human glioma cell lines in a cell line-specific manner. *PLoS One* 13(9), p. e0204159. doi: 10.1371/journal.pone.0204159

Bryan, T. M. et al. 1997. Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nat Med* 3(11), pp. 1271-1274. doi: 10.1038/nm1197-1271

Bryan, T. M. et al. 1995. Telomere elongation in immortal human cells without detectable telomerase activity. *Embo j* 14(17), pp. 4240-4248.

Burla, R. et al. 2018. Genomic instability and DNA replication defects in progeroid syndromes. *Nucleus* 9(1), pp. 368-379. doi: 10.1080/19491034.2018.1476793

Buschmann, T. et al. 2014. Enhancing the detection of barcoded reads in high throughput DNA sequencing data by controlling the false discovery rate. *BMC Bioinformatics* 15, p. 264. doi: 10.1186/1471-2105-15-264

Callen, E. and Surralles, J. 2004. Telomere dysfunction in genome instability syndromes. *Mutat Res* 567(1), pp. 85-104. doi: 10.1016/j.mrrev.2004.06.003

Capper, R. et al. 2007. The nature of telomere fusion and a definition of the critical telomere length in human cells. *Genes Dev* 21(19), pp. 2495-2508. doi: 10.1101/gad.439107

Cardoso, C. et al. 1998. Specific interaction between the XNP/ATR-X gene product and the SET domain of the human EZH2 protein. *Hum Mol Genet* 7(4), pp. 679-684. doi: 10.1093/hmg/7.4.679

Carneiro, M. O. et al. 2012. Pacific biosciences sequencing technology for genotyping and variation discovery in human data. *BMC Genomics* 13, p. 375. doi: 10.1186/1471-2164-13-375

Celli, G. B. and de Lange, T. 2005. DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. *Nat Cell Biol* 7(7), pp. 712-718. doi: 10.1038/ncb1275

Chaganti, R. S. et al. 1974. A manyfold increase in sister chromatid exchanges in Bloom's syndrome lymphocytes. *Proc Natl Acad Sci U S A* 71(11), pp. 4508-4512. doi: 10.1073/pnas.71.11.4508

Chai, W. et al. 2006. The involvement of the Mre11/Rad50/Nbs1 complex in the generation of G-overhangs at human telomeres. *EMBO Rep* 7(2), pp. 225-230. doi: 10.1038/sj.embor.7400600

Chan, C. S. and Tye, B. K. 1983. Organization of DNA sequences and replication origins at yeast telomeres. *Cell* 33(2), pp. 563-573. doi: 10.1016/0092-8674(83)90437-3

Chen, Q. et al. 2001. Two survivor pathways that allow growth in the absence of telomerase are generated by distinct telomere recombination events. *Mol Cell Biol* 21(5), pp. 1819-1827. doi: 10.1128/mcb.21.5.1819-1827.2001

Chen, Q. et al. 2016. Regulation and function of the cGAS-STING pathway of cytosolic DNA sensing. *Nat Immunol* 17(10), pp. 1142-1149. doi: 10.1038/ni.3558

Chen, X. et al. 2014. Recurrent somatic structural variations contribute to tumorigenesis in pediatric osteosarcoma. *Cell Rep* 7(1), pp. 104-112. doi: 10.1016/j.celrep.2014.03.003

Chen, Y. A. et al. 2017. Extrachromosomal telomere repeat DNA is linked to ALT development via cGAS-STING DNA sensing pathway.*Nat Struct Mol Biol*. Vol. 24. United States, pp. 1124-1131.

Cheng, C. et al. 2012. Caenorhabditis elegans POT-2 telomere protein represses a mode of alternative lengthening of telomeres with normal telomere lengths. *Proc Natl Acad Sci U S A* 109(20), pp. 7805-7810. doi: 10.1073/pnas.1119191109

Cheung, I. et al. 2004. Strain-specific telomere length revealed by single telomere length analysis in Caenorhabditis elegans. *Nucleic Acids Res* 32(11), pp. 3383-3391. doi: 10.1093/nar/gkh661

Childs, B. G. et al. 2015. Cellular senescence in aging and age-related disease: from mechanisms to therapy. *Nat Med* 21(12), pp. 1424-1435. doi: 10.1038/nm.4000

Chow, C. M. et al. 2005. Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. *EMBO Rep* 6(4), pp. 354-360. doi: 10.1038/sj.embor.7400366

Chow, T. T. et al. 2012. Early and late steps in telomere overhang processing in normal human cells: the position of the final RNA primer drives telomere shortening. *Genes Dev* 26(11), pp. 1167-1178. doi: 10.1101/gad.187211.112

Chu, H. P. et al. 2017. TERRA RNA Antagonizes ATRX and Protects Telomeres. *Cell* 170(1), pp. 86-101.e116. doi: 10.1016/j.cell.2017.06.017

Cimino-Reale, G. et al. 2001. The length of telomeric G-rich strand 3'-overhang measured by oligonucleotide ligation assay. *Nucleic Acids Res* 29(7), p. E35. doi: 10.1093/nar/29.7.e35

Cleal, K. et al. 2019. Chromothripsis during telomere crisis is independent of NHEJ, and consistent with a replicative origin. *Genome Res* 29(5), pp. 737-749. doi: 10.1101/gr.240705.118

Clynes, D. et al. 2013. The chromatin remodeller ATRX: a repeat offender in human disease. *Trends Biochem Sci* 38(9), pp. 461-466. doi: 10.1016/j.tibs.2013.06.011

Clynes, D. et al. 2015. Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX. *Nat Commun* 6, p. 7538. doi: 10.1038/ncomms8538

Clynes, D. et al. 2014. ATRX dysfunction induces replication defects in primary mouse cells. *PLoS One* 9(3), p. e92915. doi: 10.1371/journal.pone.0092915

Coluzzi, E. et al. 2017. Transient ALT activation protects human primary cells from chromosome instability induced by low chronic oxidative stress. *Sci Rep* 7, p. 43309. doi: 10.1038/srep43309

Cong, Y. S. et al. 1999. The human telomerase catalytic subunit hTERT: organization of the gene and characterization of the promoter. *Hum Mol Genet* 8(1), pp. 137-142. doi: 10.1093/hmg/8.1.137

Cong, Y. S. et al. 2002. Human telomerase and its regulation. *Microbiol Mol Biol Rev* 66(3), pp. 407-425, table of contents. doi: 10.1128/mmbr.66.3.407-425.2002

Conomos, D. et al. 2012. Variant repeats are interspersed throughout the telomeres and recruit nuclear receptors in ALT cells. *J Cell Biol* 199(6), pp. 893-906. doi: 10.1083/jcb.201207189

Cortes-Ciriano, I. et al. 2018. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *BioRxiv*, doi: https://doi.org/10.1101/333617

Costa, A. et al. 2006. Telomere maintenance mechanisms in liposarcomas: association with histologic subtypes and disease progression. *Cancer Res* 66(17), pp. 8918-8924. doi: 10.1158/0008-5472.can-06-0273

Costes, A. and Lambert, S. A. 2012. Homologous recombination as a replication fork escort: fork-protection and recovery. *Biomolecules* 3(1), pp. 39-71. doi: 10.3390/biom3010039

Counter, C. M. et al. 1992. Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *EMBO J* 11(5), pp. 1921-1929.

Court, R. et al. 2005. How the human telomeric proteins TRF1 and TRF2 recognize telomeric DNA: a view from high-resolution crystal structures. *EMBO Rep* 6(1), pp. 39-45. doi: 10.1038/sj.embor.7400314

Cox, K. E. et al. 2016. SMARCAL1 Resolves Replication Stress at ALT Telomeres. *Cell Rep* 14(5), pp. 1032-1040. doi: 10.1016/j.celrep.2016.01.011

Crabbe, L. et al. 2004. Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. *Science* 306(5703), pp. 1951-1953. doi: 10.1126/science.1103619

D'Incalci, M. and Galmarini, C. M. 2010. A review of trabectedin (ET-743): a unique mechanism of action. *Mol Cancer Ther* 9(8), pp. 2157-2163. doi: 10.1158/1535-7163.mct-10-0263

Dai, H. et al. 2009. Characterization of RNA primers synthesized by the human breast cancer cell DNA synthesome. *J Cell Biochem* 106(5), pp. 798-811. doi: 10.1002/jcb.22015

Daniali, L. et al. 2013. Telomeres shorten at equivalent rates in somatic tissues of adults. *Nat Commun* 4, p. 1597. doi: 10.1038/ncomms2602

Davis, A. J. and Chen, D. J. 2013. DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res* 2(3), pp. 130-143. doi: 10.3978/j.issn.2218-676X.2013.04.02

Davis, A. P. and Symington, L. S. 2004. RAD51-dependent break-induced replication in yeast. *Mol Cell Biol* 24(6), pp. 2344-2351. doi: 10.1128/mcb.24.6.2344-2351.2004

de Lange, T. 2004. T-loops and the origin of telomeres. *Nat Rev Mol Cell Biol* 5(4), pp. 323-329. doi: 10.1038/nrm1359

de Lange, T. 2005. Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* 19(18), pp. 2100-2110. doi: 10.1101/gad.1346005

De Vitis, M. et al. 2019. X-rays Activate Telomeric Homologous Recombination Mediated Repair in Primary Cells. *Cells* 8(7), doi: 10.3390/cells8070708

De Vitis, M. et al. 2018. Telomere Length Maintenance in Cancer: At the Crossroad between Telomerase and Alternative Lengthening of Telomeres (ALT). *Int J Mol Sci* 19(2), doi: 10.3390/ijms19020606

Deeg, K. I. et al. 2016. Cancer Cells with Alternative Lengthening of Telomeres Do Not Display a General Hypersensitivity to ATR Inhibition. *Front Oncol* 6, p. 186. doi: 10.3389/fonc.2016.00186

Deng, Y. et al. 2008. Telomere dysfunction and tumour suppression: the senescence connection. *Nat Rev Cancer* 8(6), pp. 450-458. doi: 10.1038/nrc2393

Dessain, S. K. et al. 2000. Methylation of the human telomerase gene CpG island. *Cancer Res* 60(3), pp. 537-541.

Dhayalan, A. et al. 2011. The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. *Hum Mol Genet* 20(11), pp. 2195-2203. doi: 10.1093/hmg/ddr107

Dilley, R. L. et al. 2016. Break-induced telomere synthesis underlies alternative telomere maintenance. *Nature* 539(7627), pp. 54-58. doi: 10.1038/nature20099

Dominici, M. et al. 2006. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy* 8(4), pp. 315-317. doi: 10.1080/14653240600855905

Drane, P. et al. 2010. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev* 24(12), pp. 1253-1265. doi: 10.1101/gad.566910

Draskovic, I. et al. 2009. Probing PML body function in ALT cells reveals spatiotemporal requirements for telomere recombination. *Proc Natl Acad Sci U S A* 106(37), pp. 15726-15731. doi: 10.1073/pnas.0907689106

Dugdale, H. L. and Richardson, D. S. 2018. Heritability of telomere variation: it is all about the environment! *Philos Trans R Soc Lond B Biol Sci* 373(1741), doi: 10.1098/rstb.2016.0450

Dunham, M. A. et al. 2000. Telomere maintenance by recombination in human cells. *Nat Genet* 26(4), pp. 447-450. doi: 10.1038/82586

Duquette, M. L. et al. 2004. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev* 18(13), pp. 1618-1629. doi: 10.1101/gad.1200804

Durant, S. T. 2012. Telomerase-independent paths to immortality in predictable cancer subtypes. *J Cancer* 3, pp. 67-82. doi: 10.7150/jca.3965

Duursma, A. M. et al. 2013. A role for the MRN complex in ATR activation via TOPBP1 recruitment. *Mol Cell* 50(1), pp. 116-122. doi: 10.1016/j.molcel.2013.03.006

Dynek, J. N. and Smith, S. 2004. Resolution of sister telomere association is required for progression through mitosis. *Science* 304(5667), pp. 97-100. doi: 10.1126/science.1094754

Eid, J. et al. 2009. Real-time DNA sequencing from single polymerase molecules. *Science* 323(5910), pp. 133-138. doi: 10.1126/science.1162986

Entringer, S. et al. 2011. Stress exposure in intrauterine life is associated with shorter telomere length in young adulthood. *Proc Natl Acad Sci U S A* 108(33), pp. E513-518. doi: 10.1073/pnas.1107759108

Episkopou, H. et al. 2014. Alternative Lengthening of Telomeres is characterized by reduced compaction of telomeric chromatin. *Nucleic Acids Res* 42(7), pp. 4391-4405. doi: 10.1093/nar/gku114

Faragher, R. G. et al. 1993. The gene responsible for Werner syndrome may be a cell division "counting" gene. *Proc Natl Acad Sci U S A* 90(24), pp. 12030-12034. doi: 10.1073/pnas.90.24.12030

Feng, J. et al. 1995. The RNA component of human telomerase. *Science* 269(5228), pp. 1236-1241. doi: 10.1126/science.7544491

Flynn, R. L. et al. 2015. Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors. *Science* 347(6219), pp. 273-277. doi: 10.1126/science.1257216

Frank, A. K. et al. 2015. The Shelterin TIN2 Subunit Mediates Recruitment of Telomerase to Telomeres. *PLoS Genet* 11(7), p. e1005410. doi: 10.1371/journal.pgen.1005410

Fu, Y. et al. 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* 31(9), pp. 822-826. doi: 10.1038/nbt.2623

Gan, W. et al. 2011. R-loop-mediated genomic instability is caused by impairment of replication fork progression. *Genes Dev* 25(19), pp. 2041-2056. doi: 10.1101/gad.17010011

Gibbons, R. 2006. Alpha thalassaemia-mental retardation, X linked. *Orphanet J Rare Dis* 1, p. 15. doi: 10.1186/1750-1172-1-15

Gibbons, R. J. et al. 1995. Mutations in a putative global transcriptional regulator cause X-linked mental retardation with alpha-thalassemia (ATR-X syndrome). *Cell* 80(6), pp. 837-845. doi: 10.1016/0092-8674(95)90287-2

Gibbons, R. J. et al. 1992. X-linked alpha-thalassemia/mental retardation (ATR-X) syndrome: localization to Xq12-q21.31 by X inactivation and linkage analysis. *Am J Hum Genet* 51(5), pp. 1136-1149.

Gibbons, R. J. et al. 2008. Mutations in the chromatin-associated protein ATRX. *Hum Mutat* 29(6), pp. 796-802. doi: 10.1002/humu.20734

Ginno, P. A. et al. 2012. R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell* 45(6), pp. 814-825. doi: 10.1016/j.molcel.2012.01.017

Goldberg, A. D. et al. 2010. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 140(5), pp. 678-691. doi: 10.1016/j.cell.2010.01.003

Gomes, N. M. et al. 2011. Comparative biology of mammalian telomeres: hypotheses on ancestral states and the roles of telomeres in longevity determination. *Aging Cell* 10(5), pp. 761-768. doi: 10.1111/j.1474-9726.2011.00718.x

Gordon, E. M. et al. 2016. Trabectedin for Soft Tissue Sarcoma: Current Status and Future Perspectives. *Adv Ther* 33(7), pp. 1055-1071. doi: 10.1007/s12325-016-0344-3

Graf, M. et al. 2017. Telomere Length Determines TERRA and R-Loop Regulation through the Cell Cycle. *Cell* 170(1), pp. 72-85.e14. doi: 10.1016/j.cell.2017.06.006

Grandin, N. et al. 2019. The level of activity of the alternative lengthening of telomeres correlates with patient age in IDH-mutant ATRX-loss-of-expression anaplastic astrocytomas. *Acta Neuropathol Commun* 7(1), p. 175. doi: 10.1186/s40478-019-0833-0

Green, R. A. and Kaplan, K. B. 2003. Chromosome instability in colorectal tumor cells is associated with defects in microtubule plus-end attachments caused by a dominant mutation in APC. *J Cell Biol* 163(5), pp. 949-961. doi: 10.1083/jcb.200307070

Greider, C. W. 1991. Telomeres. *Curr Opin Cell Biol* 3(3), pp. 444-451. doi: 10.1016/0955-0674(91)90072-7

Greider, C. W. and Blackburn, E. H. 1985. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* 43(2 Pt 1), pp. 405-413.

Griffith, J. D. et al. 1999. Mammalian Telomeres End in a Large Duplex Loop. *Cell* 97(4), pp. 503-514.

Grobelny, J. V. et al. 2000. ALT-associated PML bodies are present in viable cells and are enriched in cells in the G(2)/M phase of the cell cycle. *J Cell Sci* 113 Pt 24, pp. 4577-4585.

Guan, P. and Sung, W. K. 2016. Structural variation detection using next-generation sequencing data: A comparative technical review. *Methods* 102, pp. 36-49. doi: 10.1016/j.ymeth.2016.01.020

Guillou, L. and Aurias, A. 2010. Soft tissue sarcomas with complex genomic profiles. *Virchows Arch* 456(2), pp. 201-217. doi: 10.1007/s00428-009-0853-4

Haase, S. et al. 2018. Mutant ATRX: uncovering a new therapeutic target for glioma. *Expert Opin Ther Targets* 22(7), pp. 599-613. doi: 10.1080/14728222.2018.1487953

Hake, S. B. and Allis, C. D. 2006. Histone H3 variants and their potential role in indexing mammalian genomes: the "H3 barcode hypothesis". *Proc Natl Acad Sci U S A* 103(17), pp. 6428-6435. doi: 10.1073/pnas.0600803103

Hakin-Smith, V. et al. 2003. Alternative lengthening of telomeres and survival in patients with glioblastoma multiforme. *Lancet* 361(9360), pp. 836-838. doi: 10.1016/s0140-6736(03)12681-5

Harle-Bachor, C. and Boukamp, P. 1996. Telomerase activity in the regenerative basal layer of the epidermis inhuman skin and in immortal and carcinoma-derived skin keratinocytes. *Proc Natl Acad Sci U S A* 93(13), pp. 6476-6481. doi: 10.1073/pnas.93.13.6476

Harley, C. B. 2008. Telomerase and cancer therapeutics.*Nat Rev Cancer*. Vol. 8. England, pp. 167-179.

Harley, C. B. et al. 1990. Telomeres shorten during ageing of human fibroblasts. *Nature* 345(6274), pp. 458-460. doi: 10.1038/345458a0

Hayflick, L. and Moorhead, P. S. 1961. The serial cultivation of human diploid cell strains. *Exp Cell Res* 25, pp. 585-621. doi: 10.1016/0014-4827(61)90192-6

Heaphy, C. M. et al. 2011a. Altered telomeres in tumors with ATRX and DAXX mutations. *Science* 333(6041), p. 425.

Heaphy, C. M. et al. 2011b. Prevalence of the alternative lengthening of telomeres telomere maintenance mechanism in human cancer subtypes. *Am J Pathol* 179(4), pp. 1608-1615. doi: 10.1016/j.ajpath.2011.06.018

Henikoff, S. and Smith, M. M. 2015. Histone variants and epigenetics. *Cold Spring Harb Perspect Biol* 7(1), p. a019364. doi: 10.1101/cshperspect.a019364

Henson, J. D. et al. 2009. DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. *Nat Biotechnol* 27(12), pp. 1181-1185. doi: 10.1038/nbt.1587

Henson, J. D. et al. 2005. A robust assay for alternative lengthening of telomeres in tumors shows the significance of alternative lengthening of telomeres in sarcomas and astrocytomas. *Clin Cancer Res* 11(1), pp. 217-225.

Henson, J. D. et al. 2002. Alternative lengthening of telomeres in mammalian cells. *Oncogene* 21(4), pp. 598-610. doi: 10.1038/sj.onc.1205058

Hestand, M. S. et al. 2016. Polymerase specific error rates and profiles identified by single molecule sequencing. *Mutat Res* 784-785, pp. 39-45. doi: 10.1016/j.mrfmmm.2016.01.003

Heyer, W. D. et al. 2010. Regulation of homologous recombination in eukaryotes. *Annu Rev Genet* 44, pp. 113-139. doi: 10.1146/annurev-genet-051710-150955

Holliday, R. 2007. A mechanism for gene conversion in fungi. *Genet Res* 89(5-6), pp. 285-307. doi: 10.1017/s0016672308009476

Hollstein, M. et al. 1991. p53 mutations in human cancers. *Science* 253(5015), pp. 49-53. doi: 10.1126/science.1905840

Hsu, H. L. et al. 2000. Ku acts in a unique way at the mammalian telomere to prevent end joining. *Genes Dev* 14(22), pp. 2807-2812. doi: 10.1101/gad.844000

Hu, Y. et al. 2016. Switch telomerase to ALT mechanism by inducing telomeric DNA damages and dysfunction of ATRX and DAXX. *Sci Rep* 6, p. 32280. doi: 10.1038/srep32280

Huang, F. W. et al. 2013. Highly recurrent TERT promoter mutations in human melanoma. *Science* 339(6122), pp. 957-959.

Iwase, S. et al. 2011. ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. *Nat Struct Mol Biol* 18(7), pp. 769-776. doi: 10.1038/nsmb.2062

Jensen, R. B. et al. 2010. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* 467(7316), pp. 678-683. doi: 10.1038/nature09399

Jeyapalan, J. N. et al. 2008. Evidence for alternative lengthening of telomeres in liposarcomas in the absence of ALT-associated PML bodies. *Int J Cancer* 122(11), pp. 2414-2421. doi: 10.1002/ijc.23412

Jeyapalan, J. N. et al. 2005. Activation of the ALT pathway for telomere maintenance can affect other sequences in the human genome. *Human Molecular Genetics* 14(13), pp. 1785-1794. doi: 10.1093/hmg/ddi185

Ji, J. et al. 2017. Inherited germline ATRX mutation in two brothers with ATR-X syndrome and osteosarcoma. *Am J Med Genet A* 173(5), pp. 1390-1395. doi: 10.1002/ajmg.a.38184

Jiang, W. Q. et al. 2007. Identification of candidate alternative lengthening of telomeres genes by methionine restriction and RNA interference. *Oncogene* 26(32), pp. 4635-4647. doi: 10.1038/sj.onc.1210260

Jones, R. E. et al. 2014. Escape from telomere-driven crisis is DNA ligase III dependent. *Cell Rep* 8(4), pp. 1063-1076. doi: 10.1016/j.celrep.2014.07.007

Juhasz, S. et al. 2018. ATRX Promotes DNA Repair Synthesis and Sister Chromatid Exchange during Homologous Recombination. *Mol Cell* 71(1), pp. 11-24.e17. doi: 10.1016/j.molcel.2018.05.014

Jun, J. I. and Lau, L. F. 2010. Cellular senescence controls fibrosis in wound healing. *Aging (Albany NY)* 2(9), pp. 627-631. doi: 10.18632/aging.100201

Karow, J. K. et al. 2000. The Bloom's syndrome gene product promotes branch migration of holliday junctions. *Proc Natl Acad Sci U S A* 97(12), pp. 6504-6508. doi: 10.1073/pnas.100448097

Kilian, A. et al. 1997. Isolation of a candidate human telomerase catalytic subunit gene, which reveals complex splicing patterns in different cell types. *Hum Mol Genet* 6(12), pp. 2011-2019. doi: 10.1093/hmg/6.12.2011

Kim, J. et al. 2019. The macroH2A1.2 histone variant links ATRX loss to alternative telomere lengthening. *Nat Struct Mol Biol* 26(3), pp. 213-219. doi: 10.1038/s41594-019-0192-3

Kim, J. Y. et al. 2017. Alternative Lengthening of Telomeres in Primary Pancreatic Neuroendocrine Tumors Is Associated with Aggressive Clinical Behavior and Poor Survival. *Clin Cancer Res* 23(6), pp. 1598-1606. doi: 10.1158/1078-0432.ccr-16-1147

Kim, N. W. et al. 1994. Specific association of human telomerase activity with immortal cells and cancer. *Science* 266(5193), pp. 2011-2015.

Kipling, D. and Cooke, H. J. 1990. Hypervariable ultra-long telomeres in mice. *Nature* 347(6291), pp. 400-402. doi: 10.1038/347400a0

Kishi, N. and Macklis, J. D. 2004. MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions. *Mol Cell Neurosci* 27(3), pp. 306-321. doi: 10.1016/j.mcn.2004.07.006

Koschmann, C. et al. 2016. ATRX loss promotes tumor growth and impairs nonhomologous end joining DNA repair in glioma. *Sci Transl Med* 8(328), p. 328ra328. doi: 10.1126/scitranslmed.aac8228

Kourmouli, N. et al. 2005. Epigenetic regulation of mammalian pericentric heterochromatin in vivo by HP1. *Biochem Biophys Res Commun* 337(3), pp. 901-907. doi: 10.1016/j.bbrc.2005.09.132

Kratz, C. P. et al. 2017. Cancer Screening Recommendations for Individuals with Li-Fraumeni Syndrome. *Clin Cancer Res* 23(11), pp. e38-e45. doi: 10.1158/1078-0432.ccr-17-0408

Krizhanovsky, V. et al. 2008. Senescence of activated stellate cells limits liver fibrosis. *Cell* 134(4), pp. 657-667. doi: 10.1016/j.cell.2008.06.049

Kuszel, L. et al. 2015. Osteoarthritis and telomere shortening. *J Appl Genet* 56(2), pp. 169-176. doi: 10.1007/s13353-014-0251-8

La, S. H. et al. 2016. Ablation of human telomerase reverse transcriptase (hTERT) induces cellular senescence in gastric cancer through a galectin-3 dependent mechanism. *Oncotarget* 7(35), pp. 57117-57130. doi: 10.18632/oncotarget.10986

Lackner, D. H. and Karlseder, J. 2013. C. elegans survivors without telomerase. *Worm* 2(1), p. e21073. doi: 10.4161/worm.21073

Lackner, D. H. et al. 2012. Organismal propagation in the absence of a functional telomerase pathway in Caenorhabditis elegans. *Embo j* 31(8), pp. 2024-2033. doi: 10.1038/emboj.2012.61

Lafferty-Whyte, K. et al. 2009. A gene expression signature classifying telomerase and ALT immortalization reveals an hTERT regulatory network and suggests a mesenchymal stem cell origin for ALT. *Oncogene* 28(43), pp. 3765-3774. doi: 10.1038/onc.2009.238

Lai, T. P. et al. 2018. Comparison of telomere length measurement methods. *Philos Trans R Soc Lond B Biol Sci* 373(1741), doi: 10.1098/rstb.2016.0451

Lam, I. and Keeney, S. 2014. Mechanism and regulation of meiotic recombination initiation. *Cold Spring Harb Perspect Biol* 7(1), p. a016634. doi: 10.1101/cshperspect.a016634

Lamarche, B. J. et al. 2010. The MRN complex in double-strand break repair and telomere maintenance. *FEBS Lett* 584(17), pp. 3682-3695. doi: 10.1016/j.febslet.2010.07.029

Law, M. J. et al. 2010. ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. *Cell* 143(3), pp. 367-378. doi: 10.1016/j.cell.2010.09.023

Lawlor, R. T. et al. 2019. Alternative lengthening of telomeres (ALT) influences survival in soft tissue sarcomas: a systematic review with meta-analysis. *BMC Cancer* 19(1), p. 232. doi: 10.1186/s12885-019-5424-8

Lee, J. C. et al. 2015a. Alternative lengthening of telomeres and loss of ATRX are frequent events in pleomorphic and dedifferentiated liposarcomas. *Mod Pathol* 28(8), pp. 1064-1073. doi: 10.1038/modpathol.2015.67

Lee, M. et al. 2014. Telomere extension by telomerase and ALT generates variant repeats by mechanistically distinct processes. *Nucleic Acids Res* 42(3), pp. 1733-1746. doi: 10.1093/nar/gkt1117

Lee, M. et al. 2018. Telomere sequence content can be used to determine ALT activity in tumours. *Nucleic Acids Res* 46(10), pp. 4903-4918. doi: 10.1093/nar/gky297

Lee, S. S. et al. 2015b. ATM Kinase Is Required for Telomere Elongation in Mouse and Human Cells. *Cell Rep* 13(8), pp. 1623-1632. doi: 10.1016/j.celrep.2015.10.035

Lee, Y. K. et al. 2012. Prognostic value of alternative lengthening of telomeres-associated biomarkers in uterine sarcoma and uterine carcinosarcoma. *Int J Gynecol Cancer* 22(3), pp. 434-441. doi: 10.1097/IGC.0b013e31823ca017

Letsolo, B. T. et al. 2010. Fusion of short telomeres in human cells is characterized by extensive deletion and microhomology, and can result in complex rearrangements. *Nucleic Acids Res* 38(6), pp. 1841-1852. doi: 10.1093/nar/gkp1183

Levy, M. A. et al. 2008. The SWI/SNF protein ATRX co-regulates pseudoautosomal genes that have translocated to autosomes in the mouse genome. *BMC Genomics* 9, p. 468. doi: 10.1186/1471-2164-9-468

Levy, M. A. et al. 2015. ATRX promotes gene expression by facilitating transcriptional elongation through guanine-rich coding regions. *Hum Mol Genet* 24(7), pp. 1824-1835. doi: 10.1093/hmg/ddu596

Lewis, P. W. et al. 2010. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci U S A* 107(32), pp. 14075-14080. doi: 10.1073/pnas.1008850107

Li, F. et al. 2019. ATRX loss induces telomere dysfunction and necessitates induction of alternative lengthening of telomeres during human cell immortalization. *Embo j* 38(19), p. e96659. doi: 10.15252/embj.201796659

Liau, J. Y. et al. 2015a. Comprehensive screening of alternative lengthening of telomeres phenotype and loss of ATRX expression in sarcomas. *Mod Pathol* 28(12), pp. 1545-1554. doi: 10.1038/modpathol.2015.114

Liau, J. Y. et al. 2015b. Leiomyosarcoma with alternative lengthening of telomeres is associated with aggressive histologic features, loss of ATRX expression, and poor clinical outcome. *Am J Surg Pathol* 39(2), pp. 236-244. doi: 10.1097/pas.0000000000000324

Lingner, J. and Cech, T. R. 1996. Purification of telomerase from Euplotes aediculatus: requirement of a primer 3' overhang. *Proc Natl Acad Sci U S A* 93(20), pp. 10712-10717. doi: 10.1073/pnas.93.20.10712

Liu, H. et al. 2018. Telomeric Recombination Induced by DNA Damage Results in Telomere Extension and Length Heterogeneity. *Neoplasia* 20(9), pp. 905-916.

Londono-Vallejo, J. A. et al. 2004. Alternative lengthening of telomeres is characterized by high rates of telomeric exchange. *Cancer Res* 64(7), pp. 2324-2327. doi: 10.1158/0008-5472.can-03-4035

Loomis, E. W. et al. 2013. Sequencing the unsequenceable: expanded CGG-repeat alleles of the fragile X gene. *Genome Res* 23(1), pp. 121-128. doi: 10.1101/gr.141705.112

Lopez de Silanes, I. et al. 2014. Identification of TERRA locus unveils a telomere protection role through association to nearly all chromosomes. *Nat Commun* 5, p. 4723. doi: 10.1038/ncomms5723

Louis, E. J. and Borts, R. H. 1995. A complete set of marked telomeres in Saccharomyces cerevisiae for physical mapping and cloning. *Genetics* 139(1), pp. 125-136.

Lovejoy, C. A. et al. 2012. Loss of ATRX, genome instability, and an altered DNA damage response are hallmarks of the alternative lengthening of telomeres pathway. *PLoS Genet* 8(7), p. e1002772. doi: 10.1371/journal.pgen.1002772

Loyola, A. and Almouzni, G. 2007. Marking histone H3 variants: how, when and why? *Trends Biochem Sci* 32(9), pp. 425-433. doi: 10.1016/j.tibs.2007.08.004

Luger, K. et al. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* 389(6648), pp. 251-260. doi: 10.1038/38444

Lundblad, V. and Blackburn, E. H. 1993. An alternative pathway for yeast telomere maintenance rescues est1- senescence. *Cell* 73(2), pp. 347-360. doi: 10.1016/0092-8674(93)90234-h

Lundblad, V. and Szostak, J. W. 1989. A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* 57(4), pp. 633-643. doi: 10.1016/0092-8674(89)90132-3

Maes, O. C. et al. 2009. Stepwise up-regulation of microRNA expression levels from replicating to reversible and irreversible growth arrest states in WI-38 human fibroblasts. *J Cell Physiol* 221(1), pp. 109-119. doi: 10.1002/jcp.21834

Maicher, A. et al. 2012a. Deregulated telomere transcription causes replication-dependent telomere shortening and promotes cellular senescence. *Nucleic Acids Res* 40(14), pp. 6649-6659. doi: 10.1093/nar/gks358

Maicher, A. et al. 2012b. Telomeres and disease: enter TERRA. *RNA Biol* 9(6), pp. 843-849. doi: 10.4161/rna.20330

Makarov, V. L. et al. 1997. Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* 88(5), pp. 657-666. doi: 10.1016/s0092-8674(00)81908-x

Malkova, A. et al. 1996. Double-strand break repair in the absence of RAD51 in yeast: a possible role for break-induced DNA replication. *Proc Natl Acad Sci U S A* 93(14), pp. 7131-7136. doi: 10.1073/pnas.93.14.7131

Marano, D. et al. 2019. ATRX Contributes to MeCP2-Mediated Pericentric Heterochromatin Organization during Neural Differentiation. *Int J Mol Sci* 20(21), doi: 10.3390/ijms20215371

Marian, C. O. et al. 2010. The telomerase antagonist, imetelstat, efficiently targets glioblastoma tumor-initiating cells leading to decreased proliferation and tumor growth. *Clin Cancer Res* 16(1), pp. 154-163. doi: 10.1158/1078-0432.ccr-09-2850

Martin, J. A. and Buckwalter, J. A. 2001. Telomere erosion and senescence in human articular cartilage chondrocytes. *J Gerontol A Biol Sci Med Sci* 56(4), pp. B172-179. doi: 10.1093/gerona/56.4.b172

Masliah-Planchon, J. et al. 2018. Does ATRX germline variation predispose to osteosarcoma? Three additional cases of osteosarcoma in two ATR-X syndrome patients. *Eur J Hum Genet* 26(8), pp. 1217-1221. doi: 10.1038/s41431-018-0147-x

Matsuo, T. et al. 2009. Telomere-maintenance mechanisms in soft-tissue malignant fibrous histiocytomas. *J Bone Joint Surg Am* 91(4), pp. 928-937. doi: 10.2106/jbjs.g.01390

Matsuo, T. et al. 2010. Alternative lengthening of telomeres as a prognostic factor in malignant fibrous histiocytomas of bone. *Anticancer Res* 30(12), pp. 4959-4962.

Maze, I. et al. 2013. Histone regulation in the CNS: basic principles of epigenetic plasticity. *Neuropsychopharmacology* 38(1), pp. 3-22.

McDonald, K. L. et al. 2010. Presence of alternative lengthening of telomeres mechanism in patients with glioblastoma identifies a less aggressive tumor type with longer survival. *J Neuropathol Exp Neurol* 69(7), pp. 729-736. doi: 10.1097/NEN.0b013e3181e576cf

McEachern, M. J. and Blackburn, E. H. 1996. Cap-prevented recombination between terminal telomeric repeat arrays (telomere CPR) maintains telomeres in Kluyveromyces lactis lacking telomerase. *Genes Dev* 10(14), pp. 1822-1834. doi: 10.1101/gad.10.14.1822

McGraw, K. O. and Wong, S. P. 1992. A Common Language Effect Size Statistic. *Psychological Bulletin* 111(2), pp. 361-365. doi: 10.1037/0033-2909.111.2.361

Mender, I. and Shay, J. W. 2015. Telomerase Repeated Amplification Protocol (TRAP). *Bio Protoc* 5(22), doi: 10.21769/bioprotoc.1657

Mendez-Bermudez, A. et al. 2012. The roles of WRN and BLM RecQ helicases in the Alternative Lengthening of Telomeres. *Nucleic Acids Res* 40(21), pp. 10809-10820. doi: 10.1093/nar/gks862

Meselson, M. and Stahl, F. W. 1958. THE REPLICATION OF DNA IN ESCHERICHIA COLI. *Proc Natl Acad Sci U S A* 44(7), pp. 671-682. doi: 10.1073/pnas.44.7.671

Meyne, J. et al. 1990. Distribution of non-telomeric sites of the (TTAGGG)n telomeric sequence in vertebrate chromosomes. *Chromosoma* 99(1), pp. 3-10. doi: 10.1007/bf01737283

Meyne, J. et al. 1989. Conservation of the human telomere sequence (TTAGGG)n among vertebrates. *Proc Natl Acad Sci U S A* 86(18), pp. 7049-7053. doi: 10.1073/pnas.86.18.7049

Millet, C. et al. 2015. Cell populations can use aneuploidy to survive telomerase insufficiency. *Nat Commun* 6, p. 8664. doi: 10.1038/ncomms9664

Min, J. et al. 2017. Alternative Lengthening of Telomeres Mediated by Mitotic DNA Synthesis Engages Break-Induced Replication Processes. *Mol Cell Biol* 37(20), doi: 10.1128/mcb.00226-17

Mirabello, L. et al. 2015. Germline TP53 variants and susceptibility to osteosarcoma. *J Natl Cancer Inst* 107(7), doi: 10.1093/jnci/djv101

Mitson, M. et al. 2011. Functional significance of mutations in the Snf2 domain of ATRX. *Hum Mol Genet* 20(13), pp. 2603-2610. doi: 10.1093/hmg/ddr163

Mocali, A. et al. 2005. The comet assay approach to senescent human diploid fibroblasts identifies different phenotypes and clarifies relationships among nuclear size, DNA

content, and DNA damage. *J Gerontol A Biol Sci Med Sci* 60(6), pp. 695-701. doi: 10.1093/gerona/60.6.695

Moore, C. B. et al. 2010. Short hairpin RNA (shRNA): design, delivery, and assessment of gene knockdown. *Methods Mol Biol* 629, pp. 141-158. doi: 10.1007/978-1-60761-657-3_10

Morin, G. B. 1989. The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats. *Cell* 59(3), pp. 521-529. doi: 10.1016/0092-8674(89)90035-4

Morrison, S. J. et al. 1996. Telomerase activity in hematopoietic cells is associated with self-renewal potential. *Immunity* 5(3), pp. 207-216.

Motycka, T. A. et al. 2004. Physical and functional interaction between the XPF/ERCC1 endonuclease and hRad52. *J Biol Chem* 279(14), pp. 13634-13639. doi: 10.1074/jbc.M313779200

Muller, H. 1938. The Remaking of Chromosomes. Collect. Net.

Muraki, K. et al. 2012. Mechanisms of telomere loss and their consequences for chromosome instability. *Front Oncol* 2, p. 135. doi: 10.3389/fonc.2012.00135

Murnane, J. P. et al. 1994. Telomere dynamics in an immortal human cell line. *Embo j* 13(20), pp. 4953-4962.

Nabetani, A. and Ishikawa, F. 2011. Alternative lengthening of telomeres pathway: Recombination-mediated telomere maintenance mechanism in human cells. *Journal of Biochemistry* 149(1), pp. 5-14. doi: 10.1093/jb/mvq119

Nan, X. et al. 2007. Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. *Proc Natl Acad Sci U S A* 104(8), pp. 2709-2714. doi: 10.1073/pnas.0608056104

Nandakumar, J. and Cech, T. R. 2013. Finding the end: recruitment of telomerase to telomeres. *Nat Rev Mol Cell Biol* 14(2), pp. 69-82. doi: 10.1038/nrm3505

Napier, C. E. et al. 2015. ATRX represses alternative lengthening of telomeres. *Oncotarget* 6(18), pp. 16543-16558.

Nassif, N. et al. 1994. Efficient copying of nonhomologous sequences from ectopic sites via P-element-induced gap repair. *Mol Cell Biol* 14(3), pp. 1613-1625. doi: 10.1128/mcb.14.3.1613

Natarajan, S. and McEachern, M. J. 2002. Recombinational telomere elongation promoted by DNA circles. *Mol Cell Biol* 22(13), pp. 4512-4521. doi: 10.1128/mcb.22.13.4512-4521.2002

Nelson, G. et al. 2012. A senescent cell bystander effect: senescence-induced senescence. *Aging Cell* 11(2), pp. 345-349. doi: 10.1111/j.1474-9726.2012.00795.x

Nemunaitis, J. et al. 2010. A phase I study of telomerase-specific replication competent oncolytic adenovirus (telomelysin) for various solid tumors. *Mol Ther* 18(2), pp. 429-434. doi: 10.1038/mt.2009.262

Nergadze, S. G. et al. 2009. CpG-island promoters drive transcription of human telomeres. *Rna* 15(12), pp. 2186-2194. doi: 10.1261/rna.1748309

Neumann, A. A. et al. 2013. Alternative lengthening of telomeres in normal mammalian somatic cells. *Genes Dev* 27(1), pp. 18-23. doi: 10.1101/gad.205062.112

Nguyen, B. N. et al. 2009. Mechanism of dominant-negative telomerase function. *Cell Cycle* 8(19), pp. 3227-3233. doi: 10.4161/cc.8.19.9788

Nguyen, D. T. et al. 2017. The chromatin remodelling factor ATRX suppresses R-loops in transcribed telomeric repeats. *EMBO Rep* 18(6), pp. 914-928. doi: 10.15252/embr.201643078

Norris, K. et al. 2019. Telomere length predicts for outcome to FCR chemotherapy in CLL. *Leukemia* 33(8), pp. 1953-1963. doi: 10.1038/s41375-019-0389-9

Nowak-Wegrzyn, A. et al. 2004. Immunodeficiency and infections in ataxia-telangiectasia. *J Pediatr* 144(4), pp. 505-511. doi: 10.1016/j.jpeds.2003.12.046

O'Connor, M. S. et al. 2006. A critical role for TPP1 and TIN2 interaction in high-order telomeric complex assembly. *Proc Natl Acad Sci U S A* 103(32), pp. 11874-11879. doi: 10.1073/pnas.0605303103

Oganesian, L. and Karlseder, J. 2011. Mammalian 5' C-rich telomeric overhangs are a mark of recombination-dependent telomere maintenance. *Mol Cell* 42(2), pp. 224-236. doi: 10.1016/j.molcel.2011.03.015

Oganesian, L. and Karlseder, J. 2013. 5' C-rich telomeric overhangs are an outcome of rapid telomere truncation events. *DNA Repair (Amst)* 12(3), pp. 238-245. doi: 10.1016/j.dnarep.2012.12.008

Ogawa, T. and Okazaki, T. 1980. Discontinuous DNA replication. *Annu Rev Biochem* 49, pp. 421-457. doi: 10.1146/annurev.bi.49.070180.002225

Okamoto, K. et al. 2013. A two-step mechanism for TRF2-mediated chromosome-end protection. *Nature* 494(7438), pp. 502-505. doi: 10.1038/nature11873

Okuda, K. et al. 2002. Telomere length in the newborn. *Pediatr Res* 52(3), pp. 377-381. doi: 10.1203/00006450-200209000-00012

Olovnikov, A. M. 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* 41(1), pp. 181-190.

Oppel, F. et al. 2019. Loss of atrx cooperates with p53-deficiency to promote the development of sarcomas and other malignancies. *PLoS Genet* 15(4), p. e1008039. doi: 10.1371/journal.pgen.1008039

Oshima, J. 2000. The Werner syndrome protein: an update. *Bioessays* 22(10), pp. 894-901. doi: 10.1002/1521-1878(200010)22:10<894::aid-bies4>3.0.co;2-b

Painter, R. B. and Young, B. R. 1980. Radiosensitivity in ataxia-telangiectasia: a new explanation. *Proc Natl Acad Sci U S A* 77(12), pp. 7315-7317. doi: 10.1073/pnas.77.12.7315

Pan, X. et al. 2019. FANCM suppresses DNA replication stress at ALT telomeres by disrupting TERRA R-loops. *Sci Rep* 9(1), p. 19110. doi: 10.1038/s41598-019-55537-5

Pan, X. et al. 2017. FANCM, BRCA1, and BLM cooperatively resolve the replication stress at the ALT telomeres. *Proc Natl Acad Sci U S A* 114(29), pp. E5940-e5949. doi: 10.1073/pnas.1708065114

Parajuli, S. et al. 2017. Human ribonuclease H1 resolves R-loops and thereby enables progression of the DNA replication fork. *J Biol Chem* 292(37), pp. 15216-15224. doi: 10.1074/jbc.M117.787473

Parsch, D. et al. 2004. Telomere length and telomerase activity during expansion and differentiation of human mesenchymal stem cells and chondrocytes. *J Mol Med (Berl)* 82(1), pp. 49-55. doi: 10.1007/s00109-003-0506-z

Perrem, K. et al. 1999. Repression of an alternative mechanism for lengthening of telomeres in somatic cell hybrids. *Oncogene* 18(22), pp. 3383-3390. doi: 10.1038/sj.onc.1202752

Perrem, K. et al. 2001. Coexistence of alternative lengthening of telomeres and telomerase in hTERT-transfected GM847 cells. *Mol Cell Biol* 21(12), pp. 3862-3875. doi: 10.1128/mcb.21.12.3862-3875.2001

Pickett, H. A. and Reddel, R. R. 2015. Molecular mechanisms of activity and derepression of alternative lengthening of telomeres. *Nat Struct Mol Biol* 22(11), pp. 875-880. doi: 10.1038/nsmb.3106

Picketts, D. J. et al. 1996. ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. *Hum Mol Genet* 5(12), pp. 1899-1907. doi: 10.1093/hmg/5.12.1899

Picketts, D. J. et al. 1998. Comparison of the human and murine ATRX gene identifies highly conserved, functionally important domains. *Mamm Genome* 9(5), pp. 400-403. doi: 10.1007/s003359900781

Pike, A. M. et al. 2019. TIN2 Functions with TPP1/POT1 To Stimulate Telomerase Processivity. *Mol Cell Biol* 39(21), doi: 10.1128/mcb.00593-18

Pompili, L. et al. 2017. Diagnosis and treatment of ALT tumors: is Trabectedin a new therapeutic option? *J Exp Clin Cancer Res* 36(1), p. 189. doi: 10.1186/s13046-017-0657-3

Potapov, V. and Ong, J. L. 2017. Examining Sources of Error in PCR by Single-Molecule Sequencing. *PLoS One* 12(1), p. e0169774. doi: 10.1371/journal.pone.0169774

Rai, R. et al. 2016. TRF2-RAP1 is required to protect telomeres from engaging in homologous recombination-mediated deletions and fusions. *Nat Commun* 7, p. 10881. doi: 10.1038/ncomms10881

Raices, M. et al. 2008. C. elegans telomeres contain G-strand and C-strand overhangs that are bound by distinct proteins. *Cell* 132(5), pp. 745-757. doi: 10.1016/j.cell.2007.12.039

Rajagopalan, S. and Long, E. O. 2012. Cellular senescence induced by CD158d reprograms natural killer cells to promote vascular remodeling. *Proc Natl Acad Sci U S A* 109(50), pp. 20596-20601. doi: 10.1073/pnas.1208248109

Ramamoorthy, M. and Smith, S. 2015. Loss of ATRX Suppresses Resolution of Telomere Cohesion to Control Recombination in ALT Cancer Cells. *Cancer Cell* 28(3), pp. 357-369.

Ratnakumar, K. et al. 2012. ATRX-mediated chromatin association of histone variant macroH2A1 regulates alpha-globin expression. *Genes Dev* 26(5), pp. 433-438. doi: 10.1101/gad.179416.111

Reiman, A. et al. 2011. Lymphoid tumours and breast cancer in ataxia telangiectasia; substantial protective effect of residual ATM kinase activity against childhood tumours. *Br J Cancer* 105(4), pp. 586-591. doi: 10.1038/bjc.2011.266

Ren, X. et al. 2018. Alternative lengthening of telomeres phenotype and loss of ATRX expression in sarcomas. *Oncol Lett* 15(5), pp. 7489-7496. doi: 10.3892/ol.2018.8318

Rhodes, D. and Lipps, H. J. 2015. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res* 43(18), pp. 8627-8637.

Ribes-Zamora, A. et al. 2013. TRF2 interaction with Ku heterotetramerization interface gives insight into c-NHEJ prevention at human telomeres. *Cell Rep* 5(1), pp. 194-206. doi: 10.1016/j.celrep.2013.08.040

Richard, P. and Manley, J. L. 2017. R Loops and Links to Human Disease. *J Mol Biol* 429(21), pp. 3168-3180. doi: 10.1016/j.jmb.2016.08.031

Richards, E. J. and Ausubel, F. M. 1988. Isolation of a higher eukaryotic telomere from Arabidopsis thaliana. *Cell* 53(1), pp. 127-136. doi: 10.1016/0092-8674(88)90494-1

Ritchie, K. et al. 2008. Loss of ATRX leads to chromosome cohesion and congression defects. *J Cell Biol* 180(2), pp. 315-324. doi: 10.1083/jcb.200706083

Rizzo, A. et al. 2009. Stabilization of quadruplex DNA perturbs telomere replication leading to the activation of an ATR-dependent ATM signaling pathway. *Nucleic Acids Res* 37(16), pp. 5353-5364. doi: 10.1093/nar/gkp582

Roberts, R. W. and Crothers, D. M. 1992. Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. *Science* 258(5087), pp. 1463-1466. doi: 10.1126/science.1279808

Rogakou, E. P. et al. 1998. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* 273(10), pp. 5858-5868. doi: 10.1074/jbc.273.10.5858

Romano, G. H. et al. 2013. Environmental stresses disrupt telomere length homeostasis. *PLoS Genet* 9(9), p. e1003721. doi: 10.1371/journal.pgen.1003721

Roth, A. et al. 2010. Imetelstat (GRN163L)--telomerase-based cancer therapy. *Recent Results Cancer Res* 184, pp. 221-234. doi: 10.1007/978-3-642-01222-8_16

Rothblum-Oviatt, C. et al. 2016. Ataxia telangiectasia: a review. *Orphanet J Rare Dis* 11(1), p. 159. doi: 10.1186/s13023-016-0543-7

Roumelioti, F. M. et al. 2016. Alternative lengthening of human telomeres is a conservative DNA replication process with features of break-induced replication. *EMBO Rep* 17(12), pp. 1731-1737. doi: 10.15252/embr.201643169

Rudd, K. M. et al. 2007. Elevated Rates of Sister Chromatid Exchange at Chromosome Ends. *PLoS Genetics* 3(2),

Rufer, N. et al. 1999. Telomere fluorescence measurements in granulocytes and T lymphocyte subsets point to a high turnover of hematopoietic stem cells and memory T cells in early childhood. *J Exp Med* 190(2), pp. 157-167. doi: 10.1084/jem.190.2.157

Ruff, P. et al. 2016. RPA Stabilization of Single-Stranded DNA Is Critical for Break-Induced Replication. *Cell Rep* 17(12), pp. 3359-3368. doi: 10.1016/j.celrep.2016.12.003

Sager, R. 1991. Senescence As a Mode of Tumor Suppression. *Environmental Health Perspectives* 93, pp. 59-62. doi: 10.2307/3431170

Saintigny, Y. and Lopez, B. S. 2002. Homologous recombination induced by replication inhibition, is stimulated by expression of mutant p53. *Oncogene* 21(3), pp. 488-492. doi: 10.1038/sj.onc.1205040

Sakofsky, C. J. and Malkova, A. 2017. Break induced replication in eukaryotes: mechanisms, functions, and consequences. *Crit Rev Biochem Mol Biol* 52(4), pp. 395-413. doi: 10.1080/10409238.2017.1314444

Sambrock, J. et al. 1989. *Molecular Cloning, A Laboratory Manual*. Second ed. New-York: Cold Spring Harbor Laboratory Press.

Sampathi, S. and Chai, W. 2011. Telomere replication: poised but puzzling. *J Cell Mol Med* 15(1), pp. 3-13. doi: 10.1111/j.1582-4934.2010.01220.x

Sarma, K. et al. 2014. ATRX directs binding of PRC2 to Xist RNA and Polycomb targets. *Cell* 159(4), pp. 869-883. doi: 10.1016/j.cell.2014.10.019

Savic, V. et al. 2009. Formation of dynamic gamma-H2AX domains along broken DNA strands is distinctly regulated by ATM and MDC1 and dependent upon H2AX densities in chromatin. *Mol Cell* 34(3), pp. 298-310. doi: 10.1016/j.molcel.2009.04.012

Schoeftner, S. and Blasco, M. A. 2008. Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nat Cell Biol* 10(2), pp. 228-236. doi: 10.1038/ncb1685

Schwartzentruber, J. et al. 2012. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma.*Nature*. Vol. 482. England, pp. 226-231.

Seluanov, A. et al. 2007. Telomerase activity coevolves with body mass not lifespan. *Aging Cell* 6(1), pp. 45-52. doi: 10.1111/j.1474-9726.2006.00262.x

Seluanov, A. et al. 2008. Distinct tumor suppressor mechanisms evolve in rodent species that differ in size and lifespan. *Aging Cell* 7(6), pp. 813-823. doi: 10.1111/j.1474-9726.2008.00431.x

Sfeir, A. et al. 2009. Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. *Cell* 138(1), pp. 90-103. doi: 10.1016/j.cell.2009.06.021

Shampay, J. et al. 1984. DNA sequences of telomeres maintained in yeast. *Nature* 310(5973), pp. 154-157. doi: 10.1038/310154a0

Sharifi-Sanjani, M. et al. 2017. Cardiomyocyte-Specific Telomere Shortening is a Distinct Signature of Heart Failure in Humans. *J Am Heart Assoc* 6(9), doi: 10.1161/jaha.116.005086

Shay, J. W. et al. 1991. A role for both RB and p53 in the regulation of human cellular senescence. *Exp Cell Res* 196(1), pp. 33-39.

Shay, J. W. and Wright, W. E. 2006. Telomerase therapeutics for cancer: challenges and new directions. *Nat Rev Drug Discov* 5(7), pp. 577-584. doi: 10.1038/nrd2081

Singhi, A. D. et al. 2017. Alternative Lengthening of Telomeres and Loss of DAXX/ATRX Expression Predicts Metastatic Disease and Poor Survival in Patients with Pancreatic Neuroendocrine Tumors. *Clin Cancer Res* 23(2), pp. 600-609. doi: 10.1158/1078-0432.ccr-16-1113

Sjogren, C. and Nasmyth, K. 2001. Sister chromatid cohesion is required for postreplicative double-strand break repair in Saccharomyces cerevisiae. *Curr Biol* 11(12), pp. 991-995. doi: 10.1016/s0960-9822(01)00271-8

Slagboom, P. E. et al. 1994. Genetic determination of telomere size in humans: a twin study of three age groups. *Am J Hum Genet* 55(5), pp. 876-882.

Slijepcevic, P. et al. 1996. Spontaneous and radiation-induced chromosomal breakage at interstitial telomeric sites. *Chromosoma* 104(8), pp. 596-604. doi: 10.1007/bf00352299

Smith, D. J. and Whitehouse, I. 2012. Intrinsic coupling of lagging-strand synthesis to chromatin assembly. *Nature* 483(7390), pp. 434-438. doi: 10.1038/nature10895

Smith, S. et al. 1998. Tankyrase, a poly(ADP-ribose) polymerase at human telomeres. *Science* 282(5393), pp. 1484-1487. doi: 10.1126/science.282.5393.1484

Smolle, M. A. et al. 2017. A novel mutation in ATRX associated with intellectual disability, syndromic features, and osteosarcoma. *Pediatr Blood Cancer* 64(10), doi: 10.1002/pbc.26522

Sommer, A. and Royle, N. J. 2020. ALT: A Multi-Faceted Phenomenon. *Genes (Basel)* 11(2), doi: 10.3390/genes11020133

Song, J. S. et al. 2003. Adenovirus-mediated suicide gene therapy using the human telomerase catalytic subunit (hTERT) gene promoter induced apoptosis of ovarian cancer cell line. *Biosci Biotechnol Biochem* 67(11), pp. 2344-2350. doi: 10.1271/bbb.67.2344

Song, Z. et al. 2010. Lifestyle impacts on the aging-associated expression of biomarkers of DNA damage and telomere dysfunction in human blood. *Aging Cell* 9(4), pp. 607-615. doi: 10.1111/j.1474-9726.2010.00583.x

Stansel, R. M. et al. 2001. T-loop assembly in vitro involves binding of TRF2 near the 3' telomeric overhang. *Embo j* 20(19), pp. 5532-5540. doi: 10.1093/emboj/20.19.5532

Stark, J. M. et al. 2004. Genetic steps of mammalian homologous repair with distinct mutagenic consequences. *Mol Cell Biol* 24(21), pp. 9305-9316. doi: 10.1128/mcb.24.21.9305-9316.2004

Steenstrup, T. et al. 2017. Telomeres and the natural lifespan limit in humans. *Aging (Albany NY)* 9(4), pp. 1130-1142. doi: 10.18632/aging.101216

Stephens, P. J. et al. 2011. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 144(1), pp. 27-40. doi: 10.1016/j.cell.2010.11.055

Suarez, F. et al. 2015. Incidence, presentation, and prognosis of malignancies in ataxia-telangiectasia: a report from the French national registry of primary immune deficiencies. *J Clin Oncol* 33(2), pp. 202-208. doi: 10.1200/jco.2014.56.5101

Subhawong, A. P. et al. 2009. The alternative lengthening of telomeres phenotype in breast carcinoma is associated with HER-2 overexpression. *Mod Pathol* 22(11), pp. 1423-1431. doi: 10.1038/modpathol.2009.125

Sugiyama, T. and Kowalczykowski, S. C. 2002. Rad52 protein associates with replication protein A (RPA)-single-stranded DNA to accelerate Rad51-mediated displacement of RPA and presynaptic complex formation. *J Biol Chem* 277(35), pp. 31663-31672. doi: 10.1074/jbc.M203494200

Sun, L. et al. 2018. Chromatin Architectural Changes during Cellular Senescence and Aging. *Genes (Basel)* 9(4), doi: 10.3390/genes9040211

Sung, P. and Robberson, D. L. 1995. DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. *Cell* 82(3), pp. 453-461. doi: 10.1016/0092-8674(95)90434-4

Tang, J. et al. 2004a. A novel transcription regulatory complex containing death domain-associated protein and the ATR-X syndrome protein. *J Biol Chem* 279(19), pp. 20369-20377. doi: 10.1074/jbc.M401321200

Tang, P. et al. 2004b. ATRX and sex differentiation. *Trends Endocrinol Metab* 15(7), pp. 339-344. doi: 10.1016/j.tem.2004.07.006

Teng, S. C. et al. 2000. Telomerase-independent lengthening of yeast telomeres occurs by an abrupt Rad50p-dependent, Rif-inhibited recombinational process. *Molecular Cell* 6(4), pp. 947-952. doi: 10.1016/s1097-2765(05)00094-8

Teng, S. C. and Zakian, V. A. 1999. Telomere-telomere recombination is an efficient bypass pathway for telomere maintenance in Saccharomyces cerevisiae. *Mol Cell Biol* 19(12), pp. 8083-8093. doi: 10.1128/mcb.19.12.8083

Thomas, M. et al. 1976. Hybridization of RNA to double-stranded DNA: formation of R-loops. *Proc Natl Acad Sci U S A* 73(7), pp. 2294-2298. doi: 10.1073/pnas.73.7.2294

Tomaska, L. et al. 2009. Telomeric circles: universal players in telomere maintenance? *Nat Struct Mol Biol* 16(10), pp. 1010-1015. doi: 10.1038/nsmb.1660

Toupance, S. et al. 2017. Short Telomeres, but Not Telomere Attrition Rates, Are Associated With Carotid Atherosclerosis. *Hypertension* 70(2), pp. 420-425. doi: 10.1161/hypertensionaha.117.09354

Tsolou, A. et al. 2008. ssDNA fragments induce cell senescence by telomere uncapping. *Exp Gerontol* 43(10), pp. 892-899. doi: 10.1016/j.exger.2008.08.043

Ulaner, G. A. et al. 2003. Absence of a telomere maintenance mechanism as a favorable prognostic factor in patients with osteosarcoma. *Cancer Res* 63(8), pp. 1759-1763.

Uziel, T. et al. 2003. Requirement of the MRN complex for ATM activation by DNA damage. *Embo j* 22(20), pp. 5612-5621. doi: 10.1093/emboj/cdg541

van Deursen, J. M. 2014. The role of senescent cells in ageing. *Nature* 509(7501), pp. 439-446. doi: 10.1038/nature13193

Van Dyck, E. et al. 2001. Visualization of recombination intermediates produced by RAD52-mediated single-strand annealing. *EMBO Rep* 2(10), pp. 905-909. doi: 10.1093/embo-reports/kve201

van Steensel, B. et al. 1998. TRF2 protects human telomeres from end-to-end fusions. *Cell* 92(3), pp. 401-413. doi: 10.1016/s0092-8674(00)80932-0

Varley, H. et al. 2002. Molecular characterization of inter-telomere and intra-telomere mutations in human ALT cells. *Nat Genet* 30(3), pp. 301-305. doi: 10.1038/ng834

Verdun, R. E. and Karlseder, J. 2006. The DNA damage machinery and homologous recombination pathway act consecutively to protect human telomeres. *Cell* 127(4), pp. 709-720. doi: 10.1016/j.cell.2006.09.034

Villard, L. et al. 1997. Determination of the genomic structure of the XNP/ATRX gene encoding a potential zinc finger helicase. *Genomics* 43(2), pp. 149-155. doi: 10.1006/geno.1997.4793

von Morgen, P. and Maciejowski, J. 2018. The ins and outs of telomere crisis in cancer. *Genome Med* 10(1), p. 89. doi: 10.1186/s13073-018-0596-4

von Zglinicki, T. 2002. Oxidative stress shortens telomeres. *Trends Biochem Sci* 27(7), pp. 339-344. doi: 10.1016/s0968-0004(02)02110-2

von Zglinicki, T. et al. 2005. Human cell senescence as a DNA damage response. *Mech Ageing Dev* 126(1), pp. 111-117. doi: 10.1016/j.mad.2004.09.034

Voon, H. P. J. et al. 2018. Inhibition of a K9/K36 demethylase by an H3.3 point mutation found in paediatric glioblastoma. *Nat Commun* 9(1), p. 3142. doi: 10.1038/s41467-018-05607-5

Wang, X. et al. 2017. Structural basis for DAXX interaction with ATRX.*Protein Cell*. Vol. 8. Germany, pp. 767-771.

Wang, Y. et al. 2019. G-quadruplex DNA drives genomic instability and represents a targetable molecular abnormality in ATRX-deficient malignant glioma. *Nat Commun* 10(1), p. 943.

Weatherall, D. J. et al. 1981. Hemoglobin H disease and mental retardation: a new syndrome or a remarkable coincidence? *N Engl J Med* 305(11), pp. 607-612. doi: 10.1056/nejm198109103051103

Weirather, J. L. et al. 2017. Comprehensive comparison of Pacific Biosciences and Oxford Nanopore Technologies and their applications to transcriptome analysis. *F1000Res* 6, p. 100. doi: 10.12688/f1000research.10571.2

Wen, J. et al. 1998. Reconstitution of wild-type or mutant telomerase activity in telomerase-negative immortal human cells. *Hum Mol Genet* 7(7), pp. 1137-1141. doi: 10.1093/hmg/7.7.1137

Wiestler, B. et al. 2013. ATRX loss refines the classification of anaplastic gliomas and identifies a subgroup of IDH mutant astrocytic tumors with better prognosis. *Acta Neuropathol* 126(3), pp. 443-451. doi: 10.1007/s00401-013-1156-z

Wijchers, P. J. et al. 2015. Characterization and dynamics of pericentromere-associated domains in mice. *Genome Res* 25(7), pp. 958-969. doi: 10.1101/gr.186643.114

Wilhelm, T. et al. 2016. Slow Replication Fork Velocity of Homologous Recombination-Defective Cells Results from Endogenous Oxidative Stress. *PLoS Genet* 12(5), p. e1006007. doi: 10.1371/journal.pgen.1006007

Wilkie, A. O. et al. 1990. Clinical features and molecular analysis of the alpha thalassemia/mental retardation syndromes. II. Cases without detectable abnormality of the alpha globin complex. *Am J Hum Genet* 46(6), pp. 1127-1140.

Wright, W. E. et al. 1996. Telomerase activity in human germline and embryonic tissues and cells. *Dev Genet* 18(2), pp. 173-179. doi: 10.1002/(sici)1520-6408(1996)18:2<173::aid-dvg10>3.0.co;2-3

Wright, W. E. et al. 1995. Modifications of a telomeric repeat amplification protocol (TRAP) result in increased reliability, linearity and sensitivity. *Nucleic Acids Res* 23(18), pp. 3794-3795. doi: 10.1093/nar/23.18.3794

Wright, W. E. et al. 1997. Normal human chromosomes have long G-rich telomeric overhangs at one end. *Genes Dev* 11(21), pp. 2801-2809. doi: 10.1101/gad.11.21.2801

Wu, G. et al. 2003. Assembly of functional ALT-associated promyelocytic leukemia bodies requires Nijmegen Breakage Syndrome 1. *Cancer Res* 63(10), pp. 2589-2595.

Wu, G. et al. 2000. NBS1 and TRF1 colocalize at promyelocytic leukemia bodies during late S/G2 phases in immortalized telomerase-negative cells. Implication of NBS1 in alternative lengthening of telomeres. *J Biol Chem* 275(39), pp. 30618-30622. doi: 10.1074/jbc.C000390200

Wu, P. et al. 2012. Telomeric 3' overhangs derive from resection by Exo1 and Apollo and fill-in by POT1b-associated CST. *Cell* 150(1), pp. 39-52. doi: 10.1016/j.cell.2012.05.026

Wu, P. et al. 2010. Apollo contributes to G overhang maintenance and protects leading-end telomeres. *Mol Cell* 39(4), pp. 606-617. doi: 10.1016/j.molcel.2010.06.031

Wyatt, H. D. and West, S. C. 2014. Holliday junction resolvases. *Cold Spring Harb Perspect Biol* 6(9), p. a023192. doi: 10.1101/cshperspect.a023192

Wyllie, F. S. et al. 2000. Telomerase prevents the accelerated cell ageing of Werner syndrome fibroblasts. *Nat Genet* 24(1), pp. 16-17. doi: 10.1038/71630

Xiao, W. et al. 2013. Mesenchymal stem cell transformation and sarcoma genesis. *Clin Sarcoma Res* 3(1), p. 10. doi: 10.1186/2045-3329-3-10

Xue, X. et al. 2015. Functions and regulation of the multitasking FANCM family of DNA motor proteins. *Genes Dev* 29(17), pp. 1777-1788. doi: 10.1101/gad.266593.115

Xue, Y. et al. 2003. The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. *Proc Natl Acad Sci U S A* 100(19), pp. 10635-10640. doi: 10.1073/pnas.1937626100

Yang, C. Y. et al. 2015. Targeted next-generation sequencing of cancer genes identified frequent TP53 and ATRX mutations in leiomyosarcoma. *Am J Transl Res* 7(10), pp. 2072-2081.

Yeager, T. R. et al. 1999. Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Research* 59(17), pp. 4175-4179.

Yi, X. et al. 1999. Both transcriptional and posttranscriptional mechanisms regulate human telomerase template RNA levels. *Mol Cell Biol* 19(6), pp. 3989-3997. doi: 10.1128/mcb.19.6.3989

Yost, K. E. et al. 2019. Rapid and reversible suppression of ALT by DAXX in osteosarcoma cells. *Sci Rep* 9(1), p. 4544. doi: 10.1038/s41598-019-41058-8

Youngren, K. et al. 1998. Synchrony in telomere length of the human fetus. *Hum Genet* 102(6), pp. 640-643. doi: 10.1007/s004390050755

Zhan, Y. et al. 2017. Exploring the Causal Pathway From Telomere Length to Coronary Heart Disease: A Network Mendelian Randomization Study. *Circ Res* 121(3), pp. 214-219. doi: 10.1161/circresaha.116.310517

Zhang, A. et al. 2000. Frequent amplification of the telomerase reverse transcriptase gene in human tumors. *Cancer Res* 60(22), pp. 6230-6235.

Zhang, H. et al. 2017. RPA Interacts with HIRA and Regulates H3.3 Deposition at Gene Regulatory Elements in Mammalian Cells. *Mol Cell* 65(2), pp. 272-284. doi: 10.1016/j.molcel.2016.11.030

Zhang, J. M. et al. 2019a. Alternative Lengthening of Telomeres through Two Distinct Break-Induced Replication Pathways. *Cell Rep* 26(4), pp. 955-968.e953. doi: 10.1016/j.celrep.2018.12.102

Zhang, T. et al. 2019b. Strand break-induced replication fork collapse leads to C-circles, C-overhangs and telomeric recombination. *PLoS Genet* 15(2), p. e1007925. doi: 10.1371/journal.pgen.1007925

Zhong, Z. et al. 1992. A mammalian factor that binds telomeric TTAGGG repeats in vitro. *Mol Cell Biol* 12(11), pp. 4834-4843. doi: 10.1128/mcb.12.11.4834

Zhu, X. D. et al. 2000. Cell-cycle-regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres. *Nat Genet* 25(3), pp. 347-352. doi: 10.1038/77139

Zimmermann, S. et al. 2003. Lack of telomerase activity in human mesenchymal stem cells. *Leukemia* 17(6), pp. 1146-1149. doi: 10.1038/sj.leu.2402962