The role of DNA repair & regulatory proteins in the maintenance of human telomeres and their control of cellular immortalization

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Adam John Harvey

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Dr. Eric A. Hendrickson

April 2017

Acknowledgements

First, I would thank all the current and former members of the Hendrickson laboratory. It has been a wonderful environment to learn how to be a true scientist. The vast knowledge provided by my colleagues helped me with many specific problems I encountered, but more generally were just very pleasant and fun people to be around. I would especially like to thank my advisor, Dr. Eric Hendrickson. He became quite the role model for me, both as an intellectual and as a scientist. I attribute the majority of my scientific development not just to his conceptual guidance, but to his investing time, money, and training into my success. No one else will ever have impacted my intellectual development in a larger sense than he.

Thank you to my family! My parents, Brien Harvey and Jamie Just, have been a wonderful source of support, guidance, and patience. My brother, Alex Harvey, has also been a fun partner to co-mature professionally alongside. Also, he was an excellent source of video-game distractions.

Finally, thanks to my wife, Megan Harvey. She has been with me since the early days of my graduate school career and will be so for a lifetime after. Her support through this processes has been invaluable. Her sacrifices have been the foundation of my success.

Dedication

This dissertation is dedicated to the inspiration of future scientists.

The world needs a lot more science and a lot less chaos.

Abstract

Telomeres are the nucleoprotein structures that protect the ends of linear chromosomes from recognition as a double-stranded DNA break (DSB). In the absence of proper telomere function, the ends of a chromosome fuse together, creating di-centromeric chromosomes, which can no longer properly segregate at mitosis. Thus, proper telomere maintenance is absolutely essential for all eukaryotic life. Unfortunately, maintaining telomeres at a size that is protective is problematic. For example, as a consequence of "the end-replication problem," telomeres shorten incrementally during every cell cycle. These short telomeres can, in turn, function to regulate the lifespan of any given cell. Perhaps not surprisingly, therefore, humans have evolved a vast array of genes to enable telomere stability, in order to counteract any premature ageing or cell death. In order to ensure that offspring may begin their life with a default telomere length that is sufficient for stability during the organism's lifespan, stem cells must not be subjected to overall telomere shortening. Thus, all telomere shortening that a stem cell occurs during its eternal proliferation must be correspondingly compensated for by a lengthening event. This telomere elongation mechanism in essence confers cellular immortality. The most well-characterized of these cellular immortality pathways is controlled by the enzyme telomerase, which precisely elongates telomeres in a stochastic way to maintain a telomere length equilibrium. Unfortunately, this functional, essential pathway can also be conscripted to perform pathological reactions. In human cancer, all malignant growths must enable cellular immortalization to allow for their characteristic uncontrolled proliferation. In most cases this is achieved simply by the reactivation of telomerase. Interestingly, 5 to 15% of all human cancers are telomerase negative. These

cancers can be described as ALT cancers, as the Alternative Lengthening of Telomeres pathway enables their immortality. ALT, which is specific to cancer, achieves telomere elongation by aberrant recombination between telomeres. My research has found that DNA repair proteins, such as PARP1, (poly ADP ribose polymerase 1) are critical for both the maintenance of the genome and specifically for proper telomere maintenance. Furthermore, my research has demonstrated that the mutation of a single gene, ATRX, (alpha thalassemia mental retardation on the X chromosome) is an active repressor of ALT immortalization. In summary, I have contributed to the understanding of human telomere length maintenance and these studies have implications for human aging and the genesis of cancer.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Terms

Abbreviation	Full Term	Brief Description
4-OHT	4 hydroxy tamoxifen	Estrogen receptor antagonist which can be used in a CRE-ER system to induce specific recombinations
53BP1	P53 binding protein 1	End binding DSB protein that inhibits resection
6-TG	6 tioguanine	Purine analog of guanine, if metabolized by HPRT into a nucleic acid will kill the cell. Used commonly to select for HPRT null cells in gene targeting assays.
AID	Activation-induced cytidine deaminase	Involved in somatic hypermutation for antibody diversification
ALT	Alternative lengthening of telomeres	Mechanism of telomere elongation and immortalization which relies on HDR to amplify repetitive telomeric DNA
A-NHEJ	Alternative non-homologous end joining	Subpathway of NHEJ which utilizes micro-homology to facilitate the rejoining of the broken DNA
APB	ALT-associated PML body	A large macro-protein complex which commonly colocalizes with telomeres in ALT cells
ATM	Ataxia-telangiectasia mutated	A kinase which is a master regulator of the DNA damage response
ATR	Ataxia-telangiectasia and Rad3 related protein	A kinase similar to ATM, which is activated by RPA-bound ssDNA caused by resection or replication
ATRIP	ATR-interacting protein	Protein which interacts with both ATR and RPA to enable DNA checkpoint activation
ATRX	Alpha thalassemia mental retardation syndrome X-linked	ATRX is a chromatin remodeler which interacts with DAXX to facilitate H3.3 incorporation into site-specific genomic regions and is frequently mutated in ALT cancers
BER	Base excision repair	A DNA repair process to remove and replace damaged DNA bases
B-NHEJ	Backup-NHEJ	Another name for A-NHEJ
BRCA1	Breast cancer gene 1	A gene responsible for initiating and regulating the resection of DNA DSBs in HDR
BRCA2	Breast cancer gene 2	A gene which regulates the assembly of Rad51 and other necessary recombination proteins to enable HDR
Cas9	CRISPR associated protein 9	A gene from bacteria which makes endonucleic DSBs at CRISPR sites guided by an RNA molecule. Repurposed as a gene editing tool in multiple species.
ChIP	Chromatin immunoprecipitation	Technique which combines a protein immunoprecipitation with a DNA detection technique to identify protein:DNA interactions
Chk1	Checkpoint protein 1	A kinase which initiates cell cycle arrest and the DNA damage checkpoints. Usually activated by replication stress
Chk2	Checkpoint protein 2	A kinase which does is activated by ATM at the site of DNA DSBs and initiates cellular checkpoints.
C-NHEJ	Classic non-homologous end joining	DSB repair pathway which facilitates the ligation of proximal broken strands of DNA
CRE	Causes recombination	An enzyme which recombines DNA between two adjacent LoxP sites
CRISPR	Clustered regularly interspaced short palindromic repeats (CRISPR)	Technically refers to the genomic target site of a CRISPR gRNA in bacteria. Has been engineered into a gene editing platform, which uses recombinant guide

		RNAs to guide Cas9 to create breaks in DNA at specific genomic loci.
CSR	Class switch recombination	Mechanism of antibody diversification. Corresponds to the interchanging of the constant region. *e.g.* IgG->IgA
CtIP	Carboxy-terminal BRCA1 interacting protein	Endonuclease which is recruited to DNA breaks by MRN and initiates DNA resection to promote HDR
D-loop	Displacement loop	Intermediate structure in HDR caused by a 3'-ssDNA invasion into duplex DNA
DAPI	4',6-diamidino-2-phenylindole	A DNA stain that binds to the minor groove of dsDNA. Fluoresces blue.
DAXX	Death domain associated protein 6	Interacts with ATRX and is a H3.3 histone chaperone.
DNA-PKcs	DNA protein kinase catalytic subunit	A Ku-binding protein at DSBs to facilitate C-NHEJ
DN-hTERT	Dominant negative telomerase	A catalytically inactivated telomerase, which sequesters WT tert through dimerization and prevents tert activity in a cell
DSB	DNA double-strand break	A break in DNA which causes the separation of two DNA elements by breaks in both strands
dsDNA	Double-stranded DNA	The classic conformation of DNA
DT-40	Common chicken cell line	Chicken B-cell lymphoma which was oncogenically transformed with an avian virus
EJ	End joining	Another general term to describe NHEJ
FACS	Fluorescent activated cell sorting	A technique to quantify single cells according to their fluorescent signal
FBS	Fetal bovine serum	Serum taken from fetal bovines used as an additive to cell culture media to stimulate growth.
FISH	Fluorescent in situ hybridization	A molecular technique to identify the location of a specific genomic region using fluorescently labeled DNA probes.
G-overhang	Guanosine-rich 3'-overhang	3'-ssDNA overhang which is created by both the end resection problem and exonucleolytic processing at the telomere
GM847	GM clone 847	ALT cell line. SV40 transformed fibroblasts from a patient with Lesch-Nyhan syndrome
HCT116	Human colorectal tumor 116	General model cell line of choice, which was selected due to its retention of all DNA repair pathways, except mismatch repair
HDR	Homology-dependent repair	Mechanism of DNA DSB repair which requires homologous sequences for error-free repair
HPRT	Hypoxanthine-guanine phosphoribosyltransferase	Protein which converts hypoxanthine to guanine, which can be used as a selectable marker for gene targeting
HR	Homologous recombination	Commonly interchanged with HDR. Refers to the classic double-crossover HDR-subpathway
IF	Immunofluorescence	A technique which uses antibodies with florescent labels to identify sub-cellular localizations of specific proteins
IF-FISH	Immunofluorescence and Fluorescent in situ hybridization	A hybrid assay between IF and FISH. Enables the co-visualization of specific DNA sequences and proteins.
IMR90	IMR clone 90	Female lung human fibroblast cells taken from a fetus in 1977.
JFCF	Jejunal fibroblast cell line	Fibroblasts that were transformed with SV40 but not immortalized

KU70	Ku-70 kD protein	Smaller subunit of the Ku heterodimer which initiates C-NHEJ
KU86	Ku-86 kD protein	Larger subunit of the Ku heterodimer which initiates C-NHEJ
LIG1	DNA ligase 1	Ligase which is classically responsible for the sealing of Okazaki fragments in DNA replication
LIG3	DNA ligase 3	DNA ligase implicated in A-NHEJ. Also classically involved in BER
LIG4	DNA ligase 4	DNA ligase responsible for re-joining two DNA ends in C-NHEJ
LoxP	Locus of recombination in bacteriophage P1	A 34bp DNA sequence which is recognized by cre for recombination. Used in gene targeting to remove intervening DNA sequences.
MEF	Mouse embryonic fibroblast	Commonly utilized cell model generated from mouse embryos
mL3	Mitochondrial DNA ligase 3	The isoform of LIG 3 which only is present in the mitochondria, due to the mutation of downstream translation start sites, that encode the nuclear isoform.
MLS	Mitochondrial localization sequence	A protein domain which is recognized for sub-cellular localization of a given protein to the mitochondria
MMEJ	Micro-homology dependent end joining	Another synonym for A-NHEJ
MMS	Methyl methanesulfate	An alkylating agent which methylates DNA; a toxic substrate in dividing cells
MRE11	Meiotic recombination 11 protein	DNA binding protein which is primarily responsible for the initiation of HDR. A component of the MRN complex
MRN	Mre11, Nbs1, Rad50 complex	Complex that initiates end-resection and promotes HDR
NAD	Nicotinamide adenine dinucleotide	A common coenzyme which is an integral substrate in the catalysis of PAR-moieties
Nbs1	Nijmegen breakage syndrome protein 1	Another component of the MRN complex. In some species it has endonuclease activity
Neo	Neomycin resistance gene	A neomycin resistance gene which is commonly used as a drug selection marker in gene targeting
NHEJ	Non-homologous end joining	A general term for error-prone sequence-independent ligation of proximal broken DNA ends
nL3	Nuclear DNA ligase 3	The isoform of LIG 3 which only is present in the nucleus, due to the mutation of upstream MLS containing translation start sites.
ORC	Origin of replication complex	A large complex of proteins which are required to initiate DNA replication.
ORF	Open reading frame	The mature mRNA encoding DNA sequence of a gene
p16INK4a	Protein 16 kd inhibitor of cyclin dependent kinase 4a	A gene which regulates Rb and controls the initiation of the mitotic cell cycle in a G1/G0 cell.
PAR	Poly-ADP-ribose	A post-translational modification catalyzed by PARP proteins
PARP1	Poly-ADP-ribose polymerase 1	A SSB-binding protein which upon binding signals damage by synthesizing and transferring PAR onto target proteins
PAXX	Paralog of XRCC4 and XLF	An XRCC4 super-family member which participates semi-redundantly in C-NHEJ in the ligation complex.
PBZ	PAR-binding-zinc finger	A zinc-finger domain that binds to PAR
PCNA	Proliferating cellular nuclear antigen	The sliding clamp of DNA replication

PCR	Polymerase chain reaction	A molecular biology technique, where specific regions of DNA can be amplified using DNA primers to guide DNA polymerization *in vitro*. Multiple rounds produce exponential amounts of DNA
PIGA	phosphatidylinositol glycan anchor biosynthesis class A	A membrane signaling glycoprotein which functions in intracellular membrane trafficking. Can be used as a gene targeting reporter.
PML	Promyelocytic leukemia protein	Tumor suppressor gene which can conglomerate into nuclear protein complexes and which regulates genomic stability
Polβ	Polymerase *beta*	DNA polymerase recruited to replace damaged bases in BER
Polμ	Polymerase *mu*	DNA polymerase which is template independent and can add nucleotides to broken ends of DNA during C-NHEJ to provide proper ligation substrates
POT1	Protector of telomeres 1	Shelterin complex protein member which binds to telomeric ssDNA
pRb	Phosphorylated retinoblastoma	The activated form of *Rb*. A gene which is a master controller of the cell cycle initiation of a mitotic cell.
rAAV	Recombinant adeno-associated virus	A recombinant virus which is not immunogenic and can be used as either a transient gene expression or gene targeting construct
Rad50	Radiation sensitive protein 50	Third member of the MRN complex, which promotes HDR
Rad51	Radiation sensitive protein 51	ssDNA binding protein which facilities strand invasion in HDR
Rad54B	Radiation sensitive protein 54B	Binds to the head of Rad51 coated ssDNA and facilitates strand exchange
Rap1	Repressor/Activator protein 1	Shelterin complex member which facilitates protein-protein interactions
RecA	Recombination protein A	Bacterial homolog of Rad51
RNAi	RNA interference	A molecular technique to silence mRNA via the delivery of reverse complemented RNA, which will be degraded or non-functional
RPA	Replication protein A	Trimeric ssDNA binding protein complex. Though to protect ssDNA and signal its repair
SCID	Severe combined immunodeficiency	Immunodeficiency usually caused by the genetic mutation of a C-NHEJ gene
SDS	Sodium dodecyl sulfate	An amphiphilic and anionic detergent which can disrupt cellular membranes and linearize proteins
SDSA	Synthesis-dependent strand annealing	A subpathway of HDR which only has one crossover
Shelterin	Shelterin complex	The fundamental telomeric protection and regulation complex. Facilitates the recruitment of many accessory factors to the telomere
SNP	Single nucleotide polymorphisim	A variation in a single nucleotide of DNA at a specific site.
SNMI1	Sensitive to nitrogen mustard 1	Family of DNA nucleases which were discovered by sensitivity to DNA interstrand crosslinks
SSA	Single-strand assimilation	HDR subpathway initiated by the assimilation of ssDNA
SSB	Single strand DNA break	A form of DNA damage where one strand of the backbone is broken
ssDNA	Single-stranded DNA	DNA which is single stranded
STELA	Single telomere elongation assay	A linker mediated PCR technique to amplify single telomere ends to assay telomere length.

SV40	Simian virus 40	An oncogenic virus which encodes the SV40-large T antigen that can oncogenically transform healthy cells.
TBE	Tis borate EDTA	A buffer which is used in gel electrophoresis. Usually for DNA and RNA gels.
TCR	Transcription coupled repair	Form of ssDNA repair which is detected and initiated by the transcription machinery
TEL patch	Telomerase recruitment patch	A functional domain of TPP1 which facilities the recruitment of telomerase to the telomere
TERRA	Telomere RNA	The transcription product of TTAGGG telomeric DNA. It is non-coding but implicated in telomere homeostasis
TIF	Telomere induced foci	The colocalization of DNA DSB markers at a telomere
TIN2	Telomere Interacting protein 2	A shelterin complex member which facilitates protein-protein interactions
T-loop	Telomere displacement loop	A D-loop like intermediate which is caused by the G-overhang invading into the preceding repetitive DNA
TPP1	Telomere protection protein 1	Shelterin complex member which facilitates Shelterin stability and assists in the recruitment of telomere accessory factor(s) (including telomerase)
TR	Telomerase RNA component	This is a structural non-coding RNA which is bound by TERT
TRAP	Telomere repeat amplification protocol	An in vitro telomerase activity assay. Detects the *de novo* addition of telomere units on a synthetic telomere probe from cell extracts.
TERC	Telomerase RNA component	This is the gene name for TR
TERT	Telomerase	This is the reverse transcriptase which is required for normal telomere elongation
TRD	Telomere rapid deletion	The rapid loss of telomeric DNA caused by HDR-dependent resolution of the T-loop
TRF	Terminal restriction fragment	A molecular biology technique to measure telomere length by southern blot.
TRF1	Telomere repeat factor 1	DsDNA binding protein of shelterin-Very similar in structure to TRF2
TRF2	Telomere repeat factor 2	Appears to be the most essential component of Shelterin. Binds dsDNA TTAGGG
U-2 OS	Clone U-2 osteosarcoma	An ALT cell line taken from a human mesenchymal tumor.
UNG	Uracil nucleotide glycosylase	Enzyme responsible for the removal of genomic uracil
XLF	Xrcc4-like factor	Forms a filament on DNA to facilitate the rapid rejoining of broken DNA ends in C-NHEJ
XRCC1	X-ray cross complementing protein 1	Base-excision repair protein which interacts with LIG3
XRCC4	X-ray cross complementing protein 4	A cofactor for LIG4 which facilitates C-NHEJ

Chapter 1

Introduction

DNA damage, breaks, and repair

All forms of known life on Earth share the same genetic code, and many genes in the human genome can be traced all the way back in evolution to bacteria. One such incredibly well-conserved protein is Ku, a circular, doughnut shaped molecule that is responsible for DNA repair in many life-forms, and which is a subject of significant importance in **Chapter 2** of this thesis. While the specifics of Ku will be discussed later, this single example of genetic conservation throughout evolution underscores the necessity of all life to maintain and repair genetic material.

While some genes may be conserved in all species throughout evolution, Darwinian evolution by natural selection requires the existence of heritable genetic mutations. These mutations are caused by some form of DNA damage, an all-encompassing term including DNA insertions, deletions, improper covalent modifications, single-strand breaks (SSBs) and double-strand breaks (DSBs), as well as larger chromosomal aberrations, some of which will be discussed later in detail. This DNA damage originates from a variety of either exogenous or endogenous sources. Exogenous sources of DNA damage are most obvious to non-scientists, as they are often classified as carcinogens, *e.g.* ionizing radiation, UV light, heavy metals *etc*. Other sources of exogenous DNA damage can include pharmaceutically developed chemotherapies. Ironically, in the laboratory, these exogenous sources of DNA damage can all serve as tools to study DNA, and the proteins which protect it.

Endogenous sources of DNA damage are subtler in their cause, but likely contribute to the vast majority of mutations over the lifespan of an individual. For example, one recent study proposed that almost two-thirds of all human cancers are caused by endogenous random somatic mutations (Helleday *et al.* 2014). These endogenously

induced mutations could be independently detected and quantified as exogenous sources of damage tend to have unique molecular signatures, *e.g.* UV exposure causes thymidine-dimers in DNA. Although endogenous damage can be caused by (amongst other processes) oxidative DNA damage from mitochondrial dysfunction and enzymatically catalyzed DNA breaks during immune cell development, the most significant source of endogenous DNA damage is likely DNA replication, as the fidelity of DNA polymerases cannot match the large number of base pairs in a human cell, approximately 3 billion. Taken together, both exogenous and endogenous sources cause DNA damage.

Most relevant to this discussion is how specific forms of DNA damage, *i.e.* SSBs and DSBs are repaired. These two categories of DNA damage are defined by either a break in one or both strands of the DNA backbone, respectively. SSBs are the less toxic lesion *per se*, but far more common occurring approximately 10,000 to 50,000 times on a per cell, per day basis (Bradley and Kohn 1979, Vilenchik and Knudson 2003). In isolation, a SSB is relatively non-toxic, however if un-repaired a SSB can be readily converted to a DSB via DNA replication (Kuzminov 2001). These DNA DSBs, in contrast, are extremely toxic DNA lesions and mechanistically the most difficult for cells to repair (Ciccia and Elledge 2010). There are an estimated 10 spontaneous DNA DSBs per cell, per day based on metaphases studies of dividing cells (Martin *et al.* 1985, Lieber *et al.* 2003). The proper repair of these lesions is critical to maintain cellular life.

DNA SSB repair and PARP1

Given the vast incidence of SSBs in every cell, every day, a large array of SSB repair proteins have evolved to limit the potential damage of such breaks. SSBs are distinguished as being a nick in the backbone of the DNA, and not damage to the base

themselves. Although there exists a myriad of exogenous and endogenous sources of DNA damage, as long as the lesions they cause ultimately result in a SSB, they are repaired by a single common pathway (Kuzminov 2001). Indeed the cellular SSB repair pathway is mechanistically rather simple and very robust. In general, the great danger of DNA SSBs is that they can be precursors to DSBs. With that said, SSBs can pose unique challenges to the cell other than simply being readily converted to DNA. For example, SSBs can block gene transcription (Zhou and Doetsch 1993), which if unrepaired, would be equivalent to a gene-inactivating mutation. Intriguingly, various regions of the genome exhibit both variable damage and repair rates. For example, SSBs can trigger transcription coupled repair (TCR), and as such actively transcribed genomic regions tend to exhibit different repair kinetics and mutational frequencies than non-transcribed regions (Jee *et al.* 2016). Changes in the mutational rates can even be differentiated by which strand is actually transcribed, as detection of a break is much more likely to occur on the actively transcribed strand (Haradhvala *et al.* 2016). Thus SSB repair in and of its own right is a multifaceted process for preserving genomic integrity.

While there are many subpathways of SSB repair, they all share common upstream recognition proteins, including poly-ADP ribose polymerase 1 (PARP1), which binds SSBs with high affinity (Krishnakumar and Kraus 2010). This physical affinity of PARP1 for SSBs is an unbiased biochemical process, and thus *PARP1* is ubiquitously involved in all the subpathways of SSB repair. *PARP1* is one of 17 *PARP* family member genes, all which are categorized by their ability to post-translationally modify target proteins with poly-ADP-ribose (PAR) moieties. Specifically, upon the binding of a SSB, PARP1 hydrolyzes NAD^+ to generate ADP-ribose, which is then transferred by PARP1 onto a target protein (Schreiber *et al.* 2006). This reaction can be carried out sequentially to form poly-ADP-

ribose chains (hence PARP1's name), where the nascent ADP-ribose can be covalently linked to a previously transferred ADP-ribose group, rather than onto a target protein (Hassa and Hottiger 2008). These poly-ADP ribose chains can both elongate and branch out, forming large multimers consisting of 40 to 50 ADP-ribose groups. These large chains serve as signaling platforms for the recruitment of *trans*-acting response factors that mediate the subsequent repair.

Upon binding to a SSB, PARP1 also begins to modify itself with PAR moieties, which is thought to serve as a molecular scaffold for the recruitment of downstream SSB repair proteins (Gibson and Kraus 2012). The recruitment of these *trans*-acting proteins is dependent on their possession of PAR-interaction domains, which can be either large protein macrodomains, *i.e.* PARP binding zinc finger (PBZ)-domains, as well as a 8-amino acid (mostly basic residues) motif (Kleine and Luscher 2009). Canonical downstream SSB repair proteins include polymerases, such as POLβ and their accessory factors, which serve to replace any missing genetic material (Gu *et al.* 1994). Even further downstream of DNA polymerization is the ligation complex, which consists of X-ray cross complementing 1 (XRCC1) and DNA ligase 3 (LIG3), which are responsible to ligate the strands back to their proper confirmation, thus completing the reaction (Caldecott 2008).

It is important to note that not all SSB repair is the same. One notorious confusion is the difference between base excision repair (BER) and SSB repair, as they are functionally semi-redundant (Helleday 2011). BER serves to repair damage to a specific base, rather than repair the sugar backbone. BER functions by first the recognizing one of the many forms of base damage, followed by excision of the lesion to form an abasic site. This abasic site is then converted into a gap in the DNA backbone by sequential endonucleic cuts. In essence the result is two proximal SSBs, which is the source of confusing

scientific interpretation. The best example of this is the process of antibody diversification via class-switch recombination (CSR), where a cytosine deaminase, activation induced deaminase (AID), depurinates genomic cytosine causing its conversion to uracil (Muramatsu *et al.* 2000). Genomic uracil is a form of base damage, and it is excised by DNA uracil glycosylase (UNG). This causes an abasic site, which is then nicked and removed, creating a SSB. This site is certainly a suitable substrate for PARP1 binding, but the processing of this break does not require PARP1, as the pathway is a concerted reaction, and independently recruits downstream repair factors. Thus, while *PARP1* could impact the rate of this reaction in theory, it is not a BER protein.

The potential for overlapping functionality between SSB repair and BER has frequently enabled the mis-characterization of PARP1 as a BER protein. For example, PARP1 inhibitors can trap PARP1 on a transient SSB, and inhibit the canonical BER process from completion (Dantzer *et al.* 1999). Most tellingly to the importance of this distinction is the simple fact that while *PARP1*-null mice are viable (de Murcia *et al.* 1997), the loss of "true" BER proteins such as DNA polymerase beta (POLβ) (Gu et al. 1994) and XRCC1 (Tebbs *et al.* 2003) are not viable, thus suggesting that these genes serve distinct biological roles (Helleday 2011).

DNA DSB repair

DNA DSBs are extremely toxic DNA lesions and mechanistically the most difficult for cells to repair (Ciccia and Elledge 2010). There are two distinct repair pathways for DNA DSBs: homology dependent repair (HDR) and non-homologous end joining (NHEJ). The names of each pathway are self-descriptive. HDR repairs a DSB by utilizing homologous (near-identical) genetic material to "copy and paste" the correct genetic material across a

break while NHEJ facilitates the ligation of two blunt ends together, irrespective of their genetic sequence. HDR is broken down into many subpathways, including single-strand annealing (SSA), synthesis-dependent strand annealing (SDSA), and the classic double Holliday junction resolution, commonly referred to as homologous recombination (HR). There are subpathways of NHEJ as well - classic NHEJ (C-NHEJ) and alternative NHEJ (A-NHEJ). C-NHEJ is a well-conserved and rapid process which can facilitate the rejoining of adjacent ends of a DSB. A-NHEJ is hallmarked by the requirement for microhomology in the DNA sequence between the two adjacent ends to facilitate rejoining and ligation. The distinctions between these subpathways of DSB repair will be discussed below.

C-NHEJ and the Ku70/86 heterodimer

C-NHEJ is the most active DSB repair pathway in the human body. It is not significantly regulated by cell-cycle and is mechanistically simple. The canonical pathway consists of seven key genes: *Ku70*, *Ku86*, *DNA-PK*$_{cs}$, *Artemis*, *XRCC4*, *XLF4*, and *LIG4*, which are generally sufficient to rejoin most broken ends (Betermier *et al.* 2014). Two additional genes have been recently identified; one being *PAXX* (Ochi *et al.* 2015), a paralog of *XRCC4* and *XLF,* and the other being *Pol*μ (Chang *et al.* 2016), a template independent DNA polymerase. Both additional factors appear to be important to rejoin specific conformations of the DSB termini (Chang et al. 2016). Nonetheless, C-NHEJ begins with the binding of the DSB by the Ku70/86 heterodimer, an extremely abundant protein complex which has a high affinity for DNA DSBs. The mechanism of Ku binding to broken DNA ends was first elucidated by the observation that while Ku had a very high affinity for dsDNA, it required ends to facilitate its binding (Mimori and Hardin 1986). Thus

it is perhaps not surprising that the structure of this protein complex is ring-like, where the minor groove of DNA interacts with protruding, positively charged amino acids of the heterodimer, similar to the way that a nut threads onto a bolt (Walker *et al.* 2001). The Ku heterodimer is largely thought to serve as a structural signal for the recruitment of DNA-PK_{cs}, the largest known protein kinase (Lieber 2010). Interestingly, while this kinase is able to bind DNA on its own, the presence of Ku increases that affinity by 2 orders of magnitude (West *et al.* 1998). Upon DNA binding, DNA-PK_{cs} begins to phosphorylate several proteins, including itself, which enables both the activation of the DNA damage response and the recruitment of the downstream ligation complex. The signaling events include the ataxia telangiectasia mutated – checkpoint kinases 1 and 2 – and tumor suppressor protein 53 (ATM-Chk1/2-p53) activation, which signal for cell cycle to stall while the damage can be repaired. DNA-PK_{cs} also facilitates the recruitment of the downstream ligation complex via its interaction with X-ray cross-complementing 4 (XRCC4) (Costantini *et al.* 2007).

The ligation complex of C-NHEJ consists of *XRCC4*, *XLF*, *PAXX* and *LIG4*. The structure of XRCC4, XLF, and PAXX are remarkably similar, and most differ according to the relative orientation of a single C-terminal helical domain (Ochi et al. 2015). While PAXX has been shown to be essential to facilitate the rejoining of 5' incompatible ends reconstituted *in vitro* (Chang et al. 2016), *PAXX* is notably redundant only with *XLF* in living cells for VDJ recombination activity (Lescale *et al.* 2016). Thus, *PAXX* may be required to repair some rare break conformations, but it appears to be mostly functionally redundant with *XLF*. XRCC4 and XLF form a filamentous complex (presumably PAXX and XRCC4 as well), where repeated interactions via their N-terminal "head domains" can bridge the gap between the broken ends (Ropars *et al.* 2011, Andres *et al.* 2012). While

the filamentous complex structure of XRCC4 and XLF is not debated, there are semi-conflicting models of how they facilitate repair. One of the most popular is "the caterpillar model" that proposes that the two filaments wrap the DNA beginning from the Ku-bound end of the DNA, then through a larger complex-complex interaction, where the two protein-DNA filaments are attracted to one another, bringing the ends in close proximity for ligation (Reid *et al.* 2015). Regardless of which model is actually correct, LIG4 is recruited through its interaction with XRCC4 and XLF to the site of a break and can facilitate the ligation of the two broken strands (Reid et al. 2015).

Our laboratory has made several significant contributions to our current understanding of C-NHEJ. One of the initial indications that C-NHEJ might be regulated differently in humans than in other lower organisms was the finding that *Ku86* is essential in human cells (Li *et al.* 2002). In fact, the only other known organism in which Ku is essential is in the fungal organism *Ustilago maydis*, commonly known as corn smut (de Sena-Tomas et al. 2015). This essential feature in both organisms turns out to be a uniquely evolved role for Ku related to telomere maintenance, which will be discussed later (Wang *et al.* 2009). Technically, this observation required the genetic engineering of a conditionally-null human *KU86*$^{F/-}$ cell line. Upon the addition of the Cre recombinase, these cells lose their remaining functional allele of Ku and become null. Although these cells will ultimately die (*i.e.*, Ku is essential) they survive for several days. Paradoxically, an analysis of these dying *KU86*-null cells revealed that the total NHEJ activity was not at all diminished (Fattah *et al.* 2010). This was quite a surprising result, as *LIG4*$^{-/-}$ cells, have almost no detectable NHEJ activity (Fattah et al. 2010). Thus, at the time, it was reasoned that A-NHEJ activity must be compensating for the lack of C-NHEJ, and that the steady state level of total NHEJ in a *KU86*-null human cell was due to enhanced A-NHEJ. While this model turned

out to be partially true, In **Chapter 2**, I will show data that rather than A-NHEJ compensating for the total level of NHEJ is a cell, the majority of Ku-independent NHEJ is still mediated by LIG4, which implicates a novel Ku-independent direct-to-ligation mechanism by LIG4. The regulation of A-NHEJ by Ku as we had initially proposed will be discussed later in this chapter.

HDR

As aforementioned, HR is now considered a subpathway of HDR. The distinction between the HDR subpathways lies only in the downstream resolution of the homologous regions following their recognition, and thus the classic HR pathway remains the best example of HDR (Jasin and Rothstein 2013). The DSB recognition complex for HDR is the MRN complex, which consists of proteins encoded by *MRE11*, *RAD50*, and *NBS1* (Ciccia and Elledge 2010). Notably, the HDR and NHEJ recognition complexes can accumulate concurrently at the site of a DNA break (Kass and Jasin 2010), which implies some level of inter-regulatory actions between the two major pathways. The most likely explanation, is that HDR attempts to repair those breaks which C-NHEJ cannot, a decision that may involve additional factors such as the chromatin state of the DNA that is broken (Jasin and Rothstein 2013). The key step that commits the DSB repair pathway to one reliant on homology, is the 5' to 3' resection of the break, which generates recombinogenic 3'-overhangs. This resection is thought to be initiated by MRE11 and another exonuclease, CtIP, and results in the release of the Ku70/86 heterodimer (Langerak *et al.* 2011). The importance of end resection in the regulation of these two pathways is further evidenced by the effect of p53-binding protein (53BP1), which is an end-binding protein

that inhibits the ability of MRN and CtIP to initiate resection, and thus promotes NHEJ (Bunting *et al.* 2010).

Following end-resection, the resulting recombinogenic 3'-ended ssDNA is used to search for homologous sequences. First, the 3'-end is rapidly bound by replication protein A (RPA), a trimeric protein complex that binds ssDNA. RPA, in turn, recruits the ATR/ATRIP kinase complex to initiate a DNA damage response (Cimprich and Cortez 2008). In order for HDR to proceed, RPA must first be replaced by Rad51, a filamentous, single-strand binding protein that is absolutely required for the search for homologous sequences (Shinohara *et al.* 1993, San Filippo *et al.* 2008). Another accessory protein Rad54B, appears to bind to the 3' end of the recombinogenic 3' ssDNA and facilitate strand exchange (Sarai *et al.* 2006). While in many lower organisms the loss of this protein does not appear to dramatically affect HDR (Wesoly *et al.* 2006), its importance in humans was demonstrated by our laboratory via the finding that its mutation was synthetically lethal with the loss of *LIG4* (Oh *et al.* 2013). Thus, recombinogenic ssDNA, coated with Rad51 and led by Rad54B appears to be required for the search for homology in humans. This complex then facilitates the invasion and search for homologous DNA sequences, by displacing adjacent dsDNA and enabling the recombinogenic 3'-ssDNA to base pair with a cognate sequence on an adjacent chromosome. This displacement of native dsDNA by the recombinogenic ssDNA, results in the formation of a complex called a D-loop, which was originally modeled by Jack Szostak (Szostak *et al.* 1983). The mechanism of Rad51's ability to facilitate HDR is best suggested by the crystal structure of the bacterial homolog of *RAD51*, RecA, in complex with ssDNA and dsDNA (Chen *et al.* 2008). This mechanism depicts an ATP-dependent unwinding of the DNA by RecA, bound to ssDNA, thus enabling the ssDNA to hydrogen bound with cognate basepairs. Upon the recognition of

an appropriate homologous sequence, the terminus of the ssDNA-dsDNA boundary can serve as a primer to facilitate PCNA-DNA polymerase-dependent synthesis to facilitate copying of the appropriate genetic material (Sebesta *et al.* 2013).

HDR is largely regulated by the cell cycle, as recombination between sister-chromatids is vastly preferred to homologous chromosomes (Ira *et al.* 2006). This is most likely an evolutionary advantage, as the recombination with the homologous chromosome, rather than with an identical sister-chromatid, can cause a loss of heterozygosity (Moynahan and Jasin 1997). Thus, HDR tends to be limited to the S and G2 phases of the cell cycle, to ensure the presence of a proper template for such repair. This regulatory control is intimately associated with end resection and appears to be due to the regulation of 53BP1. The protection against end resection by 53BP1 requires a binding partner, called RIF-1, where together their physical interaction is stabilizing and promotes NHEJ (Chapman *et al.* 2013). This interaction is antagonized by BRCA1 and serves to remove and/or exclude 53BP1 from inhibiting end resection (Chapman *et al.* 2012). Most importantly, this exclusion of 53BP1 by BRCA1's regulation of RIF-1 is dependent on the cell cycle, and this constitutes a regulatory circuit to temporally control the occurrence of end resection, and thus HDR, limiting it only to the proper phases of the cell cycle (Escribano-Diaz *et al.* 2013).

BRCA1 and another HDR-gene *BRCA2* are both infamously well-known breast cancer susceptibility genes, where heterozygotic patients are at significant risk for breast cancer later in life (Rahman and Stratton 1998). One fascinating discovery was that *BRCA1/2* mutant patients who develop breast cancer tend to have tumors that are sensitive to PARP1 inhibition (Bryant *et al.* 2005, Farmer *et al.* 2005). The prevailing theory is that SSBs, which normally would be recognized for repair by PARP1, accumulate over time

and in a dividing cell such lesions would be converted to DSBs as a consequence of DNA replication, although other molecular models of this synthetic lethality have been proposed (Helleday 2011). Nonetheless, clinical successes of BRCA-mutant breast cancers treated with PARP1 inhibitors have been observed for some time (Fong *et al.* 2009, Tutt *et al.* 2010).

Not all *BRCA1/2* mutant breast cancers are sensitive to PARP1 inhibition (Fong et al. 2009), and the development of PARP1-inhibitor resistance has also been frequently observed. Some of the mechanisms of resistance are quite simple, such as compensatory mutations in *BRCA1/2*, which restore HDR capacity, thus eliminating the synthetic lethality (Edwards *et al.* 2008). Of more interest to this discussion, a second major mechanism to PARP1 inhibitor resistance in *BRCA1/2* mutant cancer is a compensatory mutation of 53BP1 (Lord and Ashworth 2013). This treatment resistance mechanism in cancer was originally implicated by the ability to rescue the embryonic lethality of *BRACA1*$^{-/-}$ mice (Ludwig *et al.* 1997) with a compensatory mutation in *53BP1* (Cao *et al.* 2009), which together suggests that the requirement for BRCA1 in initiating HDR is abrogated by removing 53BP1. Indeed, the first demonstration that PARP1 inhibitor resistance in *BRCA1/2*-mutant tumors might be related to the status of 53BP1 came from a study in mice, where mice that developed PARP1-inhibitor resistant *BRCA1/2* cancers had tumor-specific mutations in *53BP1* (Jaspers *et al.* 2013). The proposed mechanism for this synthetic interaction is that the absence of end protection alleviates the requirement for BRCA-dependent initiation of HR, and thus the synthetic lethality of PARP1 inhibitors in BRCA cancers is alleviated.

A-NHEJ

While HDR and C-NHEJ have been well-known and studied for decades, it has become apparent that there is an additional minor end-joining pathway present in higher eukaryotes. It has interchangeably been referred to as MMEJ (micro-homology-mediated end joining) (McVey and Lee 2008), B-NHEJ (backup-NHEJ) (Mladenov and Iliakis 2011) and A-NHEJ (alternative-NHEJ) (Nussenzweig and Nussenzweig 2007), (hereafter, A-NHEJ). The existence of A-NHEJ has actually been evident in the literature for decades (Roth and Wilson 1986), as evidence continued to mount that Ku-independent end joining, which generally required microhomologies between the sequences to facilitate its ligation, could occur (Boulton and Jackson 1996). This pathway was not explicitly defined until a kinetic analysis of reactions in the chicken DT40 cell line revealed two discrete end-joining mechanisms (Wang *et al.* 2001).

The discovery of A-NHEJ revealed that the historical accounts of Ku-independent NHEJ events had in fact been all along a discrete DNA DSB repair pathway. These reports had suggested that in the absence of Ku, end joining could occur when facilitated by microhomology (1 to 7 bp) at the termini of the break (Kabotyanski *et al.* 1998). These observations lead to the development of molecular tools to assay the frequency of A-NHEJ, using the requirement for microhomology as the definitive measure (Verkaik *et al.* 2002). Unfortunately, the protein players responsible for this pathway remained nebulous, and to this day are still debated. Adding to the complexity of the search for *bone fide* A-NHEJ proteins, is the fact that A-NHEJ is only found intermittently throughout evolution, and is particularly absent in many of the popular genetic model systems making its requirements difficult to experimentally ascertain (Simsek and Jasin 2011).

The first breakthrough with respect to identifying the proteins mediating A-NHEJ was the discovery that the *PARP1*, *XRCC1* and *LIG3* genes all appeared to be required for A-NHEJ (Audebert *et al.* 2004). This was followed by independent confirmatory reports that *LIG3* was required for A-NHEJ, as the majority of *LIG4*-independent NHEJ was abolished by knockdown of *LIG3* in mouse embryonic fibroblasts (MEFs) (Wang *et al.* 2005). While the most definitive data at the time appeared to be related to *LIG3*, *PARP1* became of significant interest perhaps in part because of its ease of study via commercially available inhibitors. PARP1's role in A-NHEJ was then elaborated by using synthetic DNA substrates with microhomology at their termini (Audebert *et al.* 2006) and it was shown that PARP1 could compete with Ku for DNA ends to facilitate A-NHEJ (Wang *et al.* 2006). Yet these reports consistently relied either on the use of pharmacological inhibition of PARP1, or genetic approaches in lower model organisms such that PARP1's role in human cells and in A-NHEJ remained equivocal. Nonetheless, the interest in A-NHEJ continued to rise as it was observed that LIG3-mediated microhomology dependent repair was causative of oncogenic translocations in mice (Simsek *et al.* 2011).

Seemingly conflicting results reported in the literature are often due to variation in the experimental model systems in which the experiments were performed. Experimental findings related to A-NHEJ are certainly one such example of abundantly conflicting literature reports. For example, one of the seminal discoveries concerning A-NHEJ was the requirement for microhomology in Ku-independent end joining in yeast (Boulton and Jackson 1996). Yet yeast only possess ancestral homologs of *LIG1* and *LIG4* (Simsek and Jasin 2011), and thus the absence of a discrete LIG3-dependent complex made this seminal finding suggestive, but not mechanistically insightful. The best example of the differences between organisms and the regulation of A-NHEJ was the finding that

oncogenic translocations are driven by A-NHEJ in mice (Simsek and Jasin 2011). This is commonly reported and microhomologies are often searched for in the discovery of new oncogenic translocations. But in humans, these interchromosomal fusions are predominently driven by C-NHEJ, and are not significantly affected by the presence or absence of LIG3 (Ghezraoui *et al.* 2014). Nonetheless, investigators continue to define human cancer translocations by the extent with which there is microhomology, although it is likely nothing more than the statistical likelihood of such events (*e.g.* there is a 1/16 chance that a translocation will have an AT-AT at is junction). My work (**Chapter 2**) indicates that A-NHEJ is detectable in the absence of LIG3, which implies that LIG3 is not solely responsible for microhomology-mediated repair. Furthermore, my work suggests that while Ku-independent A-NHEJ is LIG3-dependent, the majority of Ku-independent repair in cells remains LIG4-dependent, which I have described as “direct-ligation” repair (Oh *et al.* 2014). Furthermore, in more recent work, I show that PARP1 does not impact the cellular capacity for A-NHEJ (**Chapter 3**). Thus, while A-NHEJ remains detectable and of significant interest, its protein determinants remain uncertain.

Conceptually, while the end-to-end ligation of broken ends of DNA resembles C-NHEJ, the requirement for homology between these broken ends by A-NHEJ presumes a different mechanism than the simple joining of non-homologous ends, and rather must include some degree of end-resection to reveal micro-homologies. This process has been referred to as “micro-SSA” where resection occurs, but is not sufficient to initiate SSA (a process that requires much longer lengths of homology), but is nonetheless resected enough to enable the completion of the ligation reaction (Decottignies 2013). This model was originally based on observations from yeast, where both *RAD52* and *EXO1* (canonical SSA genes) were required for the majority of Ku-independent NHEJ (Decottignies 2007).

This was later investigated in mammalian cells (mice), which found that *CtIP* (another end resection nuclease) was required for A-NHEJ, but that this effect was actually suppressed by *RAD52* (Bennardo *et al.* 2008). The contrasting role for *RAD52* was explained away due to the fact that *RAD52* in yeast is required for almost all forms of HDR, while it appears to be HDR subpathway-specific in mice (Rijkers *et al.* 1998). Follow up work has muddied the waters even more by showing that other HDR nucleases such as MRN (Rass *et al.* 2009), as well as other HDR regulatory factors (such as *BRCA1*) (Badie *et al.* 2015), appear to be important for A-NHEJ. Although certainly not unequivocal, these reports suggested a model in mammals where a nuclease (*e.g.*, *EXO1*, *CtIP* or *MRN*) is responsible for initiating the resection required to reveal microhomologies, but the presence of downstream HDR processing factors (*e.g.*, RAD52 or BRCA1) are repressive of A-NHEJ.

Telomeres and their regulation

Telomeres are the nucleoprotein structures that protect the ends of linear chromosomes, which would otherwise be biochemically indistinguishable from a DNA DSB. The theory behind telomeres was first elegantly described by Barbara McClintock, who was performing irradiation experiments on corn cells. Notably, while studying the inheritance of genetic traits, McClintock observed that when multiple breaks occurred across many chromosomes, the chromosomes would continually rearrange over time until at some undefined point, they would stop (McClintock 1939). Later, she postulated that the chromosome ends had “healed” and noted that once end healing had occurred, that this was a permanent change (McClintock 1941). Decades later, this end healing would of course molecularly be shown to be due to the addition or generation of a telomere.

Telomere structure

While the functional property of telomeres is consistent throughout evolution, their molecular constituents are extremely species-specific. In humans, telomeric DNA consists of the hexameric repeat, 5'-TTAGGG-3' (T_2AG_3), which is repeated thousands of times to form the genetic element of a human telomere. As a consequence of its repetitive nature, the Watson-Crick strandedness of a telomere is often designated by whether it contains guanosine (TTAGGG) or cytosine (AATCCC). The double-stranded repetitive sequence extends to the distal end of the chromosome, however the very end of the chromosome is not blunt but is processed into a ssDNA 3' overhang. This overhang is often referred to as the G-overhang as it is generated by the resection of the AATCCC strand. This 3' G-overhang is inherently recombinogenic, as it is functionally similar to the DNA products created by end-resection created during HDR (Stansel *et al.* 2001). Moreover, in a fashion similar to HDR, the G-overhang is thought to invade and base pair with preceding repetitive sequences in the telomere, forming a D-loop-like telomeric loop (or "t-Loop"). In HDR such an intermediate would normally be resolved into recombinant products but at a telomere this structure appears to be relatively stable. Importantly, this re-invasion of the end of the chromosome into preceding DNA has been postulated to "hide the end of the chromosome" from DNA DSB repair-sensing proteins (Stansel et al. 2001).

The specialized sequence and structure of a human telomere is enhanced by a canonical protein complex called Shelterin. Shelterin is a 6-member protein complex, which forms a proteinaceous cap over the end of the chromosome and which regulates telomere dynamics (de Lange 2010). The six shelterin complex proteins are:

- TRF1 and TRF2, which bind to the dsDNA T_2AG_3
- POT1, which binds to the ssDNA T_2AG_3
- Rap1, TPP1, and TIN2, which mediate protein:protein interactions but do not bind DNA themselves

As a whole, the Shelterin complex coordinates telomere dynamics and stability. In just one example, separation-of-function mutants in the basic domain in the N-terminus of TRF2 cause HDR-dependent resolution of the t-loop, and the entire telomere is eliminated from the chromosome in the form of a circle (Wang *et al.* 2004).

While the fundamental stability of a telomere depends upon how Shelterin interacts with various components of the DNA repair and replication machinery, there are many levels of telomere dynamics. For instance, while Shelterin is only a 6-member protein complex in humans, there are over 300 reported Shelterin-interacting proteins (Lee *et al.* 2011). These "telomere accessory factors" are also very important for the regulation of telomeric structure. One critical role of these factors is to generate the G-overhang, which is thought to be mediated by Apollo, an SNM1 (sensitive to nitrogen mustard 1) family nuclease related to the C-NHEJ nuclease, Artemis (Wu *et al.* 2010). The absence of Apollo results in a dramatic loss of the G-overhang, and is causative of the leading (G-rich) end to fuse with other chromosomes (Wu et al. 2010). Importantly, Apollo is recruited to the telomere transiently though an interaction with TRF2 (van Overbeek and de Lange 2006). This interaction has been crystalized, and determined to occur at a canonical "docking site" in TRF2, where proteins with a TRF2-interacting motif, Y-X-L-X-P, can be recruited to telomeres to perform essential functions, without themselves having specificity

for telomeric DNA (Chen *et al.* 2008). Thus, telomere accessory factors recruited by Shelterin can perform essential telomere roles.

In addition to TRF2, many other Shelterin proteins regulate telomere accessory factors. For example, TPP1 was classically thought to be relegated to mediating Shelterin:Shelterin interactions (Liu *et al.* 2004), and was known to regulate telomere length (Keegan *et al.* 2005). Later, it was discovered that a key TPP1 regulatory domain, the TEL patch, is required for telomerase recruitment to the telomere (Nandakumar *et al.* 2012). Patient mutations in the TPP1 TEL patch result in inherited diseases of bone marrow failure (Guo *et al.* 2014). This implied that while telomerase is a very well-studied telomere protein (discussed later) it is nonetheless a "telomere accessory protein," as its functional utility requires Shelterin components to facilitate its recruitment.

The absence of a functional telomere, either the natural erosion of telomeric DNA by "the end replication problem" (discussed later), or by the removal of Shelterin, results in phenotypes that resemble the (mis)repair of DSBs. Indeed, eroded or uncapped telomeres are magnets for a bevy of DSB repair factors. The visualization of such proteins with fluorescently-labeled antibodies in metaphase cells results in foci at the telomeres called telomere induced foci (TIFs). For example, the natural erosion of telomeric DNA that ultimately triggers cellular senescence is tightly correlated with the appearance of TIFs (Sedivy 1998). Thus, a telomere's essential function is preventing the end of a chromosome from being recognized as a DNA DSB.

One of the deleterious outcomes of attracting DSB repair proteins to a telomere is that fusions and chromosomal translocations can ensue. For instance, the absence of TRF2 bound to telomeric DNA permits end:end C-NHEJ-mediated inter-chromosomal fusions, and is causative of the transformation of independent linear chromosomes into one

"snake-like" mega chromosome (Celli and de Lange 2005). This phenotype, in turn, is dependent on the DNA damage signaling and DSB repair components of a cell. For example, the telomere fusions in a *TRF2*$^{-/-}$ cell can be rescued by a compensatory mutation in *ATM* (Denchi and de Lange 2007). Similarly, the *TRF2*$^{-/-}$ fusion events can also be rescued by the loss of C-NHEJ components, such as *LIG4* (Celli and de Lange 2005) or *Ku70* (Celli *et al.* 2006). Interestingly, in *TRF2*$^{-/-}$:*Ku70*$^{-/-}$ double mutants not only is the end-to-end fusion activity rescued (as would be predicted by Ku's role in C-NHEJ), but the telomeres also exhibited high levels of sister-chromatid exchanges, which is a hallmark of HDR and a phenotype not observed in the *TRF2*$^{-/-}$ mutants (Celli et al. 2006). Altogether, these data suggest that the absence of Shelterin can induce telomeric fusions that are mediated by C-NHEJ. Moreover, in the accompanying absence of C-NHEJ, or more specifically Ku, HDR can also act upon the telomeres.

The above data suggest that Ku has additional roles at a telomere rather than simply facilitating the C-NHEJ-mediated fusing of dysfunctional telomeres. What that role(s) is, however, has been difficult to decipher using model systems. Thus, in *Saccharomyces cerevisiae*, *KU*70-null strains exhibit telomere shortening and telomere dysfunction (Boulton and Jackson 1998). However, in *Arabidopsis thaliana*, the loss of Ku results in massive telomere elongation (Riha *et al.* 2006, Boltz *et al.* 2014) whereas the loss of Ku in DT40 chicken cells has no apparent telomere phenotype (Wei *et al.* 2002). To add even more confusion to this conundrum is the fact that two different laboratories reported that the loss of Ku in mice resulted in either no effect on telomeres (Samper *et al.* 2000) or significant telomere shortening (d'Adda di Fagagna *et al.* 2001). Thus, the use of model systems has, in the case of Ku and telomere biology, shed no light whatsoever on what Ku might be doing in a human cell. On a superficial level, it is not unreasonable to imagine

that Ku could be a telomere protein since it is a DNA end-binding protein and is responsible for the protection of those ends, which would be a key feature of a telomeric protein. Generally, however, Ku exists to facilitate the rejoining of DNA ends, which is antithetical to the purpose of a telomeric protein. One potential solution to this paradox is evidence that Ku binding to DNA ends is generally repressive of other forms of DNA DSB repair (Wang et al. 2006, Fattah et al. 2010, Langerak et al. 2011). The implication of these observations would be that Ku bound onto or associated with a telomere protects the telomere from other forms of DNA repair, but it itself is not sufficient to initiate C-NHEJ of telomeres. This hypothesis is consistent with genetic work from other species.

Ku is essential for cellular life in only 2 known species: *Ustilago maydis* (corn smut) (de Sena-Tomas et al. 2015) and *Homo sapiens* (Wang et al. 2009). In both cases, the essential role of Ku appears to repress an HDR-dependent telomere rapid deletion (TRD). In corn smut, the loss of Ku is causative of a massive G_2/M arrest, which is likely due to the DNA damage signaling initiated from uncapped telomeres (de Sena-Tomas et al. 2015). This arrest, and the TRD, are both able to be rescued by the deletion of *MRE11* (de Sena-Tomas et al. 2015), a key HDR gene. In humans, *KU86*-null cells also exhibit TIFs and TRD which is coincident with the formation of extra-chromosomal circular telomeric DNA molecules, called T-circles, which is the product of HDR-dependent resolution of the telomere (Wang et al. 2004, Wang et al. 2009). Thus, it seems as if Ku can be a telomere factor in some species by directly binding onto the telomeric DNA, but that in human cells (where Ku's DNA binding role has been usurped by Shelterin) Ku's role lies predominately in excluding HDR factors from pathologically accessing the telomere.

In the absence of Ku in human cells many of the chromosomes do not fuse end-to-end (i.e., a chromosomal translocation); rather, adjacent sister chromatids fuse end to end as an intra-chromosomal fusion. While inter-chromosomal fusions are generally regarded as lethal events, intra-chromosomal fusions are causative of prolonged chromosomal instability, but are nonetheless viable (Murnane 2010). These fusions are causative of "breakage-fusion-bridging" (BFBs) cycles where, during mitosis as the chromosomes are segregated, a chromosome that has a sister-sister fusion would behave as one large di-centric chromosome (Murnane 2010). This process was originally observed and named by Barbara McClintock (McClintock 1939). Classically it was thought that the BFBs were caused by the molecular forces of the opposing spindles. However, over decades of research, evidence has mounted against this model. One key observation was that the mitotic spindle was estimated to only be able to generate 0.5 to 1.5 nN, whereas a mitotic chromosome has been estimated to withstand up to 100 nN of force (Houchmandzadeh *et al.* 1997). Other evidence from *Saccharomyces cerevisiae* indicates that mitotic chromosomes can withstand the spindle force, but are broken by the mitotic process (Haber *et al.* 1984, Lopez *et al.* 2015). In human cells, dicentric chromosomes can persist from the beginning of anaphase and extend even beyond cytokinesis (Maciejowski *et al.* 2015). Thus, these bridges can persist for several hours following mitosis, and are enclosed by the pre-existing nuclear envelope, which connects the two nuclei of the cells (Maciejowski et al. 2015). Ultimately, the nuclear exonuclease TREX1 gains access to this bridge, coverts the dsDNA to ssDNA, thus weakening the strand considerably, such that torsional strain can ultimately break the connection (Maciejowski et al. 2015). The discovery of this mechanism also explains the observation that BFBs can be causative of a localized hypermutation event called chromothripsis, that tends to occur at the point of

a BFB, as a significant amount of genetic material is likely to be damaged by this process (Li *et al.* 2014, Maciejowski et al. 2015). These BFBs also play a critical role in regulating the transition of a pre-malignant mortal cell into an immortalized cancerous one. Thus, Ku's role in human telomere biology appears be to act as a deterrent to HDR and when Ku fails to perform this function, intra-chromosomal fusions and BFB cycling result, which is ultimately toxic to the cell.

Telomeres, telomerase, and cellular immortality

As researchers developed our understanding of the semi-conservative nature of DNA replication, the observation that DNA polymerases replicated DNA in only one orientation (5'>3') implicated that there could be an "end replication problem." That is, as the DNA replication fork approaches the end of the chromosome the leading strand can be completely replicated, whereas the requirement for an RNA primer to initiate lagging strand DNA synthesis implied that the terminal nucleotides of the lagging strand could not be replicated (Olovnikov 1973). This conceptual realization had profound implications; *i.e.* as cells continued to divide *in absentia* of any other regulatory processes, linear chromosomes would gradually erode ends-in over consecutive cellular generations. This theory matched the empirical observations of Leonard Hayflick, who reported that fibroblast cells seemed to senesce in culture after 40 to 60 population doublings, thus infamously refuting the theory that all cells are immortal (Hayflick and Moorhead 1961). Later it was determined that this was indeed due to telomere erosion generating TIFs and a potent DNA damaging signal that causes cell cycle arrest (Feldser and Greider 2007). Furthermore, the realization of the "end replication problem" implicated that there must be a hitherto unknown post-replicative process capable of chromosomal end re-elongation.

The molecular determinants of these profound ideas would later be confirmed and are discussed below.

In the summer of 2015, I had the privilege of meeting the Nobel Prize winner Dr. Elizabeth Blackburn, (awarded for her work on cellular immortalization), with whom I spent several hours into the "wee hours of the morning," as she put it. In 1985, Dr. Blackburn, and her graduate student at the time, Carol Greider, were able to detect the laddering of a radiolabeled telomeric DNA probe in cell-free extracts of the single-celled eukaryotic organism *Tetrahymena thermopilia* (Greider and Blackburn 1985). They noted that the telomeric probe was incorporated in a ladder-like fashion, such that each telomeric probe was incorporated as an elongating string of DNA in specific length intervals. Shortly following this discovery, the enzymatic components of this reaction were identified in yeast, the gene names of which were aptly named for the consequence of their mutation: Ever Shorter Telomeres-1, 2, and 3 (Lundblad and Szostak 1989). This enzyme was later discovered to also exist in human cancer cells, and ultimately become known to be called telomerase (Morin 1989).

It is now well accepted that the maintenance of telomere length by telomerase is required for cellular life. In animals, the expression of telomerase is responsible for maintaining the telomeric content of dividing stem cells to ensure that subsequent generations do not lose telomere length over time. Strangely, the absence of telomerase in mice does not cause any dramatic initial phenotype (Blasco *et al.* 1997); a paradox later explained by the extremely long telomeres of in-bred laboratory strains of mice. Nonetheless, these mice exhibited telomere erosion at approximately 5 kb (~10% of their original telomere length) every generation. By the fourth generation, the mice were lethargic, sick, and not fertile: attributes confirming the critical function of telomerase to

preserving genomic integrity over time (Blasco et al. 1997). Notably, these phenotypes would not take several generations to manifest in humans, as mice have telomeres approximately 10X longer than humans (de Lange *et al.* 1990), and have substantially shorter intergenerational time periods of about 6 months. In addition to germ cells, telomerase expression is critical to maintain other proliferative cell types. Most notably, the immune system requires clonal expansion of antigen-activated cells during the stimulation of the adaptive immune response. This rapid clonal expansion can cause telomere shortening, even in the presence of telomerase (Epel *et al.* 2004). As a consequence, immune deficiency is a classic phenotype associated with mutations of telomere maintenance genes in mammals. Thus, telomere length maintenance by telomerase is a critical component of many physiological functions.

Given the importance of telomerase to human life it is not surprising that the complex has been extensively studied. Human telomerase consists of two major subunits: the protein component, TERT, and the RNA subunit, TR (Martinez and Blasco 2011). TERT is a large, multi-domain well-conserved reverse transcriptase. Although TERT has been the subject of intense study, the full-length protein has not been crystalized to date. Thus, our structural information relies on the crystallization of various domains of the protein, including the interface of the protein with the RNA component (Qiao and Cech 2008). Another domain of TERT, the TEN domain, partially interacts with telomeric DNA, but this binding interface still allows the TR access to the end of the chromosome (Lue 2005, Lue and Li 2007). An additional domain in TERT facilitates its interaction with the TEL patch of TPP1, which also enables the recruitment of telomerase to the telomere (Nandakumar et al. 2012), as has been discussed earlier in this section. In addition to TERT's

interactions with telomeric DNA and Shelterin, TERT also makes an essential interaction with TR that are required telomere elongation.

TR has both structural and functional components that enable the formation of a functional telomerase complex. TR has several RNA secondary structures, which form loops and hairpins that are critical for TR to interact with both TERT and telomeric DNA. Two such loops are especially critical for the interaction of TR and TERT, those being the CR4/5 domain (a series of two proximal looped secondary structures) (Bley *et al.* 2011), and the pseudoknot domain (Qiao and Cech 2008). In addition to these structural components, TR is also functionally critical for telomerase's reverse transcription, as it both primes and extends off the terminal end of the G-overhang and provides the template for reverse transcription (Nandakumar and Cech 2013) (I accidentally confirmed this requirement by wrongfully adding RNAseA to the buffer of several *in vitro* telomerase assays, which were blank). Thus, TERT and TR are both equally critical for telomerase activity.

De novo telomere elongation is a processive enzymatic reaction, the extent of which is generally defined by how many repeats are synthesized prior to the complex's dissociation from the telomere (Greider 1991). The extent of telomerase's enzymatic processivity varies wildly between species such that, for example, human telomerase is much more processive than murine telomerase (Prowse *et al.* 1993). The processivity of telomerase is likely determined by the TR structure, rather than the enzyme itself, as simple mutations in TR can dramatically impact the processivity of telomerase across multiple species (Chen and Greider 2003). In addition to TR being a regulator of processivity, the Shelterin proteins TPP1 and POT1 have also been identified as regulating this effect of telomerase (Wang *et al.* 2007). Lastly, the relative abundance of

telomeric ends, compared to the amount of telomerase complexes also affects processivity; the fewer complexes there are relative to the amount of chromosomal ends results in a more processive enzyme (Wang et al. 2007). Taken together, these observations imply that the degree of telomerase processivity is likely a stochastic effect dependent on the abundance of the protein itself, protein:protein interactions and the number of telomeres present (the last feature of which is relevant to inter-species comparisons and polyploid cancers).

The dynamic interplay between constitutive telomere shortening due to the end replication problem, and telomere re-elongation by telomerase regulates the proliferative capacity for normal cellular life. The evolution of this cellular-lifespan regulatory network is likely attributable to it being a strong mechanism to prevent cancer. According to the somatic mutation theory of cancer, a dividing cell that acquires an oncogenic mutation enters into a Darwinian evolution scheme, which selects for faster and more invasive cell types. In a feed-forward mechanism, accelerated cellular growth rates are known to be contributors of additional genomic mutations, which is referred to as oncogenic-replicative stress (Negrini *et al.* 2010). Additional mutations in these pre-cancerous cells can inactivate various checkpoints which would otherwise provide limitations on cell growth. Cells that have been oncogenically activated and have lost various cellular checkpoints through mutation can be considered to be "transformed." The history of this term originates from virology gene transfer, as infection by oncogenic viruses was one of the first-discovered direct means of changing healthy cells into cancerous ones (Eddy *et al.* 1962). Notably, in these instances the virus encodes proteins, such as the SV (simian virus) 40 large-T antigen, which can simultaneously induce cell growth and inhibit checkpoints (Goldstein 1990). However, in spite of the historical nomenclature regarding

oncogenic transformation, the majority of cancers become transformed by non-viral means (Negrini et al. 2010).

The proliferation of transformed cells is ultimately limited by telomere shortening, where in the absence of normal cellular checkpoints, cells will enter telomere crisis. Telomere crisis is a term used to describe the critically short length of a telomere, where the telomere has eroded to a non-functional extent. The kinetics of telomere erosion that leads to crisis in human fibroblasts can be determined using a technique called Single-Telomere Analysis (STELA), which is a linker-mediated PCR assay using a chromosome-specific subtelomeric PCR forward primer, and the end of the chromosome as a reverse primer to amplify single telomeric ends (Baird *et al.* 2003). Thus, the process of telomere erosion can be precisely monitored as cells age and enter telomere crisis. At the point of telomere crisis, most cells die, due to the formation of the previously discussed BFBs and chromosomal fusions, most of which are lethal.

An escape from telomere crisis is required for a transformed cell to become a *bona fide* malignancy, and represents one fundamental "Hallmark of Cancer" (Hanahan and Weinberg 2011). Although this escape rarely occurs, Darwinian evolution strongly selects those few cells capable of enabling a telomere elongation mechanism (TEM). There are two known TEMs, 1) the reactivation of telomerase or 2) the initiation of the Alternative Lengthening of Telomeres (ALT) pathway. Telomerase reactivation (through gene amplification, activating promoter mutations, *etc.*) occurs in ~85% of cases, while the ALT pathway is less common. The telomerase TEM in cancer is straightforward; the restoration of telomerase activity in cancer cells enables them to not be subjected to telomere shortening, thus achieving cellular immortality. One striking example of this effect is that the first human cell line to ever be grown continuously in culture, HeLa cells

(which express telomerase), were taken from a woman, Henrietta Lacks, in the 1950's. Today, Ms. Lacks would be one of the oldest living (but very likely deceased) people on Earth and yet, HeLa cells are robustly still growing in culture to this day.

In contrast to telomerase, ALT activation is not common although it is enriched in several subtypes of cancers. It is most often found in malignancies of mesenchymal origin, such as osteosarcomas (Henson and Reddel 2010). For reference, I have graphically depicted the process of a malignant cell's telomere shortening, evasion of cell checkpoints, and ultimate escape from telomere crisis through the use of either telomerase or ALT (**Figure 1.1**). The mechanism of ALT and its genetic regulators will be discussed below, and in several subsequent chapters of this dissertation.

ALTernative immortalization, telomeric chromatin, and ATRX

While ALT only occurs in about 15% of human cancers, it remains an attractive avenue for developing new chemotherapies. This is because ALT is a tumor-specific molecular mechanism for telomere elongation, unlike telomerase which is commonly utilized by both cancerous cells and healthy (*e.g.*, stem cells) ones. ALT facilitates telomere elongation via a DNA recombination reaction, akin to a telomeric "copy-and-paste" function (Cesare and Reddel 2010). As telomeres are all identical in their DNA sequence content, a short telomere can utilize another telomere as a template for HDR-dependent elongation. This allows for cells to elongate all of their telomeres, irrespective of an individual telomere's length, as the reaction is solely dependent on the sequence homology. Thus, while telomerase is a tightly regulated enzymatic function, no such regulation is present for ALT cells, and as a consequence individual telomere length can vary wildly in ALT cells,

whereas telomerase-positive cells tend to stochastically achieve a relatively consistent balance in telomere length.

Heterogenous telomere length is one of the many hallmarks of an ALT cell (Bryan *et al.* 1995). The discovery that *TRF2* was repressive of HDR-dependent TRD inspired the finding that another hallmark of ALT cells is the presence of extra-chromosomal T-circles (Wang et al. 2004). These circular telomeric DNA molecules have been proposed to serve as templates for rolling-circle amplification, where a telomeric end using HDR could invade the T-circle and then elongate indefinitely around the circular template. Other hallmarks of ALT include "ALT-associated PML bodies" (APBs) (Yeager *et al.* 1999). APBs are large multimeric DNA:protein complexes, and are proposed to be the site of active telomeric recombination within the nucleus (Conomos *et al.* 2014). One final hallmark of ALT requires experimental manipulation, where the genetic insertion of an exogenous DNA element into telomeric DNA will be "copied and pasted" into other telomeres on distinct chromosomes (Dunham *et al.* 2000). While, these hallmarks of ALT are mechanistically understood, the molecular control regarding the initiation of ALT remained somewhat of a mystery until the discovery that a gene encoding a chromatin remodeler, α-Thalassemia Mental Retardation Syndrome X-linked gene (*ATRX*), was commonly mutated in ALT cancers (Heaphy *et al.* 2011). The molecular mechanism of ATRX's regulation of ALT is likely a complicated one, as ATRX's role as a telomeric chromatin remodeler does not have a direct effect on telomere length itself.

The role of ATRX in the regulation of ALT requires an understanding of the chromatin state of telomeres. Telomeres are generally regarded as heterochromatic, as they are largely structural DNA elements that to do not encode any genes. The importance of this heterochromatic state is exemplified by the requirement of *SIRT6* for proper telomere

homeostasis. Sirt6 is a histone 3 lysine 9 (H3K9) deacetylase, which removes acetyl groups from histone H3 and causes chromatin compaction (Michishita *et al.* 2008). The loss of *SIRT6* in mice is causative of a dramatic premature ageing phenotype and telomere dysfunction, presumably because of the mice's inability to properly condense the telomeres into a protective state (Michishita et al. 2008). Thus, the ability of telomeres to compact as heterochromatin is critical to their protective function.

Although telomeres are heterochromatic, they can be transcribed into non-coding, telomeric repeat RNA (TERRA) (Azzalin *et al.* 2007). TERRA, in turn, regulates telomeric chromatin through an interaction with TRF1 and TRF2 that functions to maintain telomeric H3K9me3 and to recruit the origin of replication (ORC) to the telomere for proper DNA replication (Deng *et al.* 2009). As the expression of any RNA is dependent on the chromatin state, TERRA acts in a feedback function, where a more euchromatic telomere favors TERRA production, which promotes its transition back to heterochromatin (Eid *et al.* 2015). Thus, TERRA can be used as a reporter of how compact the telomere is in each cell, where higher cellular TERRA-levels imply less of a heterochromatic state.

While the heterochromatic state of a telomere is important, there are more chromatin regulators than just *SIRT6* or TERRA. Indeed, as aforementioned, the mutation of the chromatin remodeler, ATRX, is associated with the genesis of ALT. Recent work has uncovered a role for ATRX at the telomere, where it interacts with the H3.3 chaperone, DAXX, to facilitate the deposition of H3.3 into telomeric DNA. This regulatory network was first implicated by a series of CHIP-seq experiments in MEFs, which identified that the histone variant H3.3 was enriched in telomeric DNA (Goldberg *et al.* 2010). In addition, distinct H3.3 chaperone complexes were localized to different genomic regions. Notably, while depletion of the H3.3 chaperone, Hira, dramatically affected intragenic levels of

H3.3, the depletion of DAXX/ATRX was causative of a loss in telomeric H3.3 (Goldberg et al. 2010). The functional consequence of this was identified soon after, as the loss of *ATRX* in human embryonic stem cells is causative of telomere destabilization (Wong *et al.* 2010). With respect to ALT, the major impact of this telomeric H3.3 regulatory network was realized when a subset of cancers lacking *ATRX* or *DAXX*, all appeared to have all enabled their immortality utilizing the ALT pathway (Heaphy et al. 2011). Upon a more thorough examination of a large array of ALT cancer lines, it was revealed that nearly 100% of ALT cell lines were either missing, or mutated in *ATRX* or *DAXX* (Lovejoy *et al.* 2012), although this association was much stronger for *ATRX* than for *DAXX*. Initial attempts to establish a causality between ATRX mutation and ALT were performed by cell hybridization experiments, which showed that *ATRX* mutations segregated with the ALT phenotype (Bower *et al.* 2012). These seminal reports from 2010-2012 spawned a new field of investigation into the role of ATRX in regulating and/or preventing the genesis of ALT, including my work presented in **Chapters 4 and 5**. How the presence or absence of a chromatin remodeling complex for H3.3 was regulating ALT, remained mechanistically completely unclear.

While ATRX is a chromatin remodeling protein responsible for H3.3 incorporation, the role of H3.3 in inhibiting ALT was first demonstrated by mutation of another H3.3. chaperone, *ASF1* (O'Sullivan *et al.* 2014). ASF1 is histone chaperone that assists with HIRA-mediated replication-independent H3.3 deposition into chromatin (Abascal *et al.* 2013). Mutation of *ASF1* in telomerase-positive cells resulted in a dramatic induction of ALT-activity (O'Sullivan et al. 2014). Interestingly, this phenotype required the cells to have long telomeres, which suggests that one reason human telomere length is so tightly regulated is that they are inherently recombinogenic. As ALT-activation is a mechanism

by which cells can escape telomere crisis, the requirement for long telomeres and the lack of any obvious patient-mutations in *ASF1* make this finding mechanistically insightful, but unlikely to be relevant for human disease. One hypothesis arising from this mechanistic insight was that the chromatin state of the telomere may regulate its propensity to be accessed by different DNA repair pathways, such that the absence of telomeric H3.3 was somehow enabling HDR at the telomere. This idea was founded based on the frequent observation that the chromatin-state of a DNA end can affect DNA DSB pathway choice (Miller *et al.* 2010). Thus, the prevailing theory is that ATRX-mediated H3.3 incorporation into telomeric DNA is repressive of telomeric recombination (O'Sullivan and Almouzni 2014) and that the loss of this repression permits ALT.

Given that ablation of *ASF1* was sufficient to induce ALT in telomerase-positive cells, the consequences of an *ATRX* mutation is notably different. In **Chapter 4**, I will show that genetic mutation of *ATRX* in telomerase-positive cells does not induce ALT. The lack of ALT activity in these cells is especially interesting, as the re-expression of *ATRX* in ALT cells does repress ALT activity (Clynes *et al.* 2015). Thus, it appears that ATRX represses ALT, but the true persistence of ALT requires additional modifications or processes to enable an ALT cell. In **Chapter 5**, I confirm this hypothesis by showing that the absence of ATRX significantly enhances the frequency of ALT immortalization and, at a molecular level, enables telomeric recombinations at critically shortened telomeres.

Rationale

Telomeres are essential for eukaryotic cellular life. Their fundamental evolutionary purpose is to protect the ends of linear chromosomes from being recognized as a DNA DSB. Intriguingly, multicellular organisms appear to have evolved a programmed cell

death capacity, by intentionally not elongating their telomeres. In humans, this process represents one of the strongest tumor-suppressing mechanisms, as virtually all malignancies must enable cellular immortality to facilitate limitless growth. The interplay between telomeres and DNA DSB repair is ironically further intertwined, as not only does the DSB repair pathway that recognizes a dysfunctional telomere dramatically influence that cell's likelihood of achieving immortality, but 15% of all cancers (ALT^+) appear to require a subpathway of DSB repair to elongate their telomeres and remain immortalized. Thus, the work I present in the following chapters is intended to clarify the regulation of pathway choice within DSB repair, the role of DNA break sensing proteins in protecting normal telomeres, and the role of *ATRX* in suppressing the genesis of ALT.

Figures

Figure 1.1

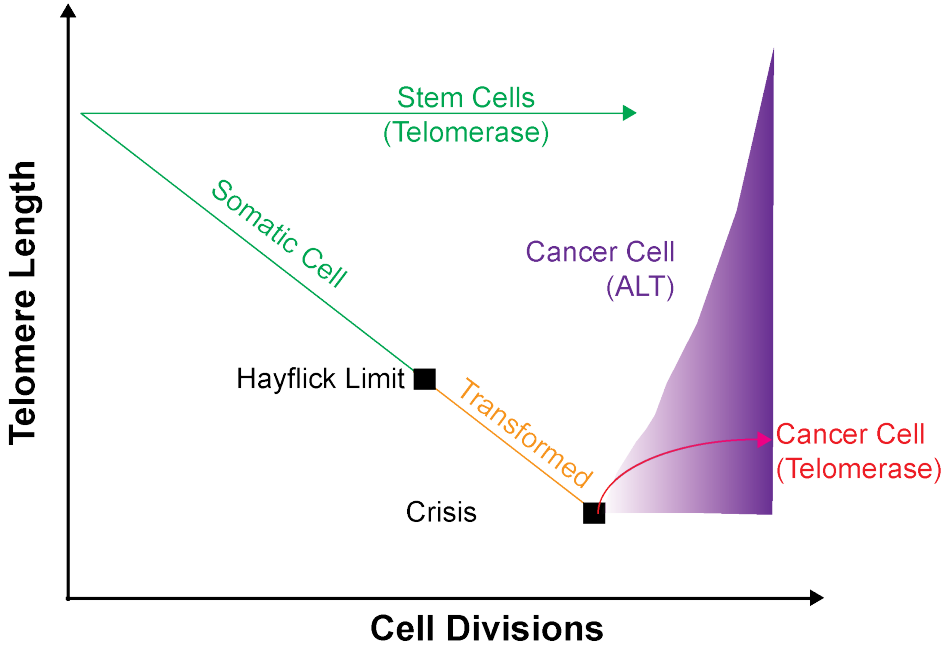


Figure 1.1 Graphical representation of telomere attrition and immortalization.

The relationship between telomere length and number of cellular divisions is plotted. Healthy cells either maintain their telomeres with telomerase (*e.g.* stem cells), or gradually lose their telomeres due to the end replication problem. As somatic cells continue to age and lose telomeres, they will at some point initiate cellular checkpoints and stop growing, referred to as "The Hayflick Limit." Cells that are transformed may continue to divide until they enter telomere crisis, at which point telomeres no longer function and the chromosomes undergo BFB. At this juncture most cells will die. Some rare survivors may initiate either telomerase re-expression or ALT to enable cellular immortality and become malignant.

Chapter 2

DNA Ligase 3 and DNA Ligase 4 carry out genetically distinct forms of end joining in human cells

[Published in Oh, S., A. Harvey, J. Zimbric, Y. Wang, T. Nguyen, P. J. Jackson and E. A. Hendrickson (2014). *DNA Repair* 21: 97-110. PMID: 24837021]

Author Contributions:

- Oh, Sehyun, Zimbric.Y., and Wang, Y. constructed all cell lines used, except for Ku:LIG3. All experiments describing single mutant cells (Figs. 2.1 to 2.7) were performed by Oh., S.
- Harvey, A. was responsible for the experimental design and execution of all experiments related to the double-mutant cell line characterization as well as the proposed model explaining all results discovered (Figs. 2.8 to 2.12), with assistance from Nguyen, T. This also included working on the construction of Ku:LIG3 double mutant.
- Jackson, P.J. completed construction of the Ku86:LIG3 line.

Synopsis

Ku-dependent C-NHEJ is the primary DNA EJing repair pathway in mammals. Recently, an additional EJing repair pathway (A-NHEJ) has been described. Currently, the mechanism of A-NHEJ is obscure although a dependency on LIG3 is often implicated. To test the requirement for LIG3 in A-NHEJ we constructed a *LIG3* conditionally-null human cell line using gene targeting. Nuclear EJing activity appeared unaffected by a deficiency in *LIG3* as, surprisingly, so were random gene targeting integration events. In contrast, *LIG3* was required for mitochondrial function and this defined the gene's essential activity. Human *KU86*:*LIG3* and *KU86*:*LIG4* double knockout cell lines, however, demonstrated that LIG3 is required for the enhanced A-NHEJ activity that is observed in *KU86*-deficient cells. Most unexpectedly, however, the majority of EJing events remained LIG4-dependent. In conclusion, although human LIG3 has an essential function in mitochondrial maintenance, it is dispensable for most types of nuclear DSB repair, except for the A-NHEJ events that are normally suppressed by Ku. Moreover, we describe that a robust Ku-independent, LIG4-dependent repair pathway exists in human somatic cells.

Introduction

A serious challenge to genomic integrity is the occurrence of a DNA DSB (Lieber 2010). To avoid the pathological outcomes that result from even a single unrepaired DNA DSB, all cells have developed efficient DSB repair pathways. In most organisms, there are two major pathways: HDR and C-NHEJ (Hartlerode and Scully 2009, Kass and Jasin 2010). HDR is preferentially used in lower organisms, however in mammals — and particularly in human cells — the majority of DSBs are repaired via C-NHEJ.

C-NHEJ facilitates the direct ligation of the broken ends of a DSB. Since the DNA termini formed at DSBs are, however, often complex and can contain non-ligatable end groups, the repair of such DNA lesions may require the processing of the ends prior to ligation (Lieber 2010, Boboila *et al.* 2012). This requirement often leads to the loss or addition of nucleotides from either side of the DSB, making C-NHEJ "error-prone". The mechanism of C-NHEJ-mediated DSB repair postulates that Ku binds to the DSB ends, where it recruits downstream C-NHEJ factors that facilitate processing (Yano *et al.* 2011). Finally LIG4, in association with XRCC4 and XLF, performs the end ligation reaction (Lieber 2010). This linear, stepwise model for C-NHEJ may be oversimplified as there is evidence that LIG4, XRCC4 and XLF may perform roles both upstream and downstream in the repair process (Andres *et al.* 2012, Roy *et al.* 2012, Cottarel *et al.* 2013).

There is an additional EJing pathway present in higher eukaryotes. It has interchangeably been referred to as MMEJ (McVey and Lee 2008), B-NHEJ (Mladenov and Iliakis 2011) and A-NHEJ (Nussenzweig and Nussenzweig 2007), (hereafter, A-NHEJ). Unlike the HDR and C-NHEJ pathways, which are conserved from bacteria to man, the A-NHEJ pathway has evolved in a somewhat checkered manner and can only

be detected in about a third of eukaryotic genomes (Simsek and Jasin 2011). It is presumed that an end-binding factor besides Ku is required to bind onto the broken DNA ends, stabilize them, protect them from random nuclease degradation and finally funnel the ends into the A-NHEJ pathway (Wang *et al.* 2006). Then, because microhomology is frequently used to mediate the repair event, some end resection is required (Zha *et al.* 2009). Alignment activities to bring the microhomologies into register are also needed, followed by the action of a flap-like nuclease to trim non-base paired regions and finally a ligation complex to covalently link the ends back together (Gottlich *et al.* 1998). Because the pathway uses microhomology to mediate the repair event, deletions always accompany the repair event, as does loss of one of the blocks of microhomology (Boboila et al. 2012).

Several laboratories have made dedicated attempts to identify A-NHEJ factors. In particular, a brute-force nuclear extract fractionation protocol identified LIG3 (Simsek and Jasin 2011), heretofore known only for its role in BER (base excision repair), as the candidate ligase required for A-NHEJ (Wang *et al.* 2005). Using guilt-by-association as a scientific rationale, PARP1 and XRCC1, two proteins known to interact with LIG3 during BER, were subsequently identified as also being involved in A-NHEJ (Audebert *et al.* 2004, Wang et al. 2006, Cheng *et al.* 2011). PARP1 is presumed to compete with Ku for binding to broken DNA ends thereby dictating pathway choice (Wang et al. 2006, Cheng et al. 2011) whereas XRCC1 appears to act as a chaperone for LIG3 (Caldecott *et al.* 1994). Additional factors have also been implicated in A-NHEJ. Thus, CtIP and the MRN complex — factors known to be involved in the end resection events required for HR — have also been implicated in the end resection steps of A-NHEJ (Rass *et al.* 2009, Xie *et al.* 2009, Zhuang *et al.* 2009, You and Bailis 2010, Zhang and Jasin 2011).

If the factors needed for A-NHEJ are not completely defined and the A-NHEJ reaction mechanism nebulous, it is also fair to say that the biological role(s) of A-NHEJ is even more poorly understood. Most of the current interest in A-NHEJ, however, stems from its implicated use in the chromosomal translocations that are present in cancer cells. Sequencing of human cancer genomes has revealed that many (Bentley *et al.* 2004, Mattarucchi *et al.* 2008, Tsai *et al.* 2008), albeit not all (Bunting and Nussenzweig 2013) chromosomal translocations have microhomology at their breakpoint junctions, which implicates A-NHEJ in their genesis. This hypothesis gained support from work in which LIG3 conditionally-null murine cells showed decreased translocation frequencies and reduced microhomology usage (Simsek and Jasin 2010, Gao *et al.* 2011, Simsek *et al.* 2011, Simsek *et al.* 2011). An additional biological process where A-NHEJ has been implicated is in the random insertion events associated with rAAV (recombinant adeno-associated virus)-mediated gene targeting. Although rAAV can mediate high frequency gene targeting, the majority of the viral integration events still occur randomly (Khan *et al.* 2011). Moreover, our laboratory has reported that a reduction in the C-NHEJ proteins Ku70 (Fattah *et al.* 2008) and LIG4 (Oh *et al.* 2013) had almost no impact on the random rAAV integration rate — implying that these events may be mediated instead by A-NHEJ. In summary, although A-NHEJ was a neglected subject for many years, in the past decade it has proven itself to be an increasingly interesting and biologically relevant topic.

A key feature of A-NHEJ is its dependence on LIG3 (Wang et al. 2005). Unlike the other ligases, LIG3 is molecularly heterogeneous (Lakshmipathy and Campbell 1999, Perez-Jannotti *et al.* 2001, Simsek and Jasin 2011). Thus, alternative translation initiation generates mitochondrial and nuclear forms of LIG3, which either contain or lack a MLS (mitochondrial localizing sequence), respectively (Lakshmipathy and Campbell 1999).

The existence of LIG3 isoforms implies diverse functional roles for LIG3. One experimental approach to unraveling the complexity of LIG3 is to generate a *LIG3*-deficient model system, which has already been accomplished in the chicken cell line, DT40 (Arakawa *et al.* 2012), and in the mouse (Puebla-Osorio *et al.* 2006, Gao et al. 2011, Simsek et al. 2011). In these systems, the gene is essential due to its presumed requirement for mitochondrial DNA replication. Moreover, in *LIG3* conditionally-null mice no obvious nuclear DNA repair phenotypes could be detected (Gao et al. 2011, Simsek et al. 2011). The extrapolation of these studies to humans is unfulfilled as neither LIG3 patients nor LIG3-deficient human cell systems have been described.

In this study, we conditionally inactivated the *LIG3* gene in the HCT116 human cell line and confirmed that the loss of *LIG3* results in death due to mitochondrial dysfunction. We also constructed a cell line that exclusively expressed a mitochondrial-only form of LIG3. The nuclear LIG3 deficiency in this cell line caused a growth retardation, but it did not affect the overall NHEJ repair activity nor did it result in hypersensitivity to DNA damaging agents. Unexpectedly, we also demonstrate that LIG3-dependent A-NHEJ does not mediate rAAV random integration events. These findings were extended by constructing human cell lines that were doubly-null in either *KU86* and *LIG3* or *KU86* and *LIG4*. These experiments demonstrated that LIG3 is required for the enhanced A-NHEJ activity that is observed in Ku-deficient cells and that the vast majority of repair events in a *KU86*-null cell are still LIG4-dependent. In conclusion, human LIG3 has an essential function in mitochondrial maintenance, however it is dispensable for most nuclear DSB repair, except for the A-NHEJ that is normally suppressed by Ku. In addition, we demonstrate that human cells have a robust Ku-independent, but LIG4-dependent EJing activity.

Material and methods

PCR primers

The sequences for all primers referenced in the Materials and Methods are listed in **Table 2.1**.

Construction of *LIG3*-null cells

Conditional and non-conditional knockout *LIG3* targeting vectors were constructed as described, with a few modifications (Kohli *et al.* 2004, Khan et al. 2011). For the conditional knockout vector, the left and right homology arms, the latter of which contained the floxed exon 5, were generated by PCR. For the left homology arm, Exon5_LARM_F1 and Exon5_LARM_SacR1 primers were used. For the right homology arm, the PCR products generated from Exon5_RARM_XhoF1 x Exon5_RARM_R1 primers and Exon5_KpnF1 x Exon5_XhoR1 primers were ligated after XhoI restriction enzyme digestion. The relevant homology arms and a NEO (neomycin-resistance) gene cassette were assembled together through ligations followed by unique restriction enzyme digestions, and then cloned into the pAAV-MCS vector. A non-conditional knockout targeting vector was generated in a similar way, but it did not include the floxed exon 5 sequences. To select for productively infected cells, the rAAV-infected cells were incubated in 1 mg/ml G418-containing media for approximately 2 weeks. At this time, genomic DNA was purified from all G418-resistant clones and PCR was used to screen for the subset of those in which correct targeting had taken place. Targeted clones were screened with Exon5_SC_F2 x NeoR2 primers, and retargeted clones were confirmed by LIG3_LArm_F3 x LIG3_RArm_R2 primers.

Construction of a *KU86$^{flox/-}$*:*LIG3$^{mtio/-}$* cell line

Conditional and knock-in *LIG3* targeting vectors were constructed as described above. In the first round of targeting, *KU86$^{flox/-}$* cells (Wang et al. 2009) were infected with a conditional knockout (i.e., "flox:NEO") vector, as described in the creation of the *LIG3$^{flox/-}$* cell line. After confirming correct targeting, clones were Cre-treated, and screened for the loss of the drug selection cassette (NEO), but retention of the conditional (floxed) exon, thus yielding a *KU86$^{flox/-}$*:*LIG3$^{+/flox}$* cell line. In the second round of targeting, these cells were infected with a rAAV knock-in vector, which introduced ATC point mutations into the two closely spaced ATGs that enable nuclear LIG3 expression. For the rAAV *LIG3* knock-in targeting vector construction, the left homology arm was PCR-amplified from wild-type HCT116 genomic DNA using LIG3Mut_LArmF and LIG3Mut_LArmR primers. The right homology arm was amplified using LIG3Mut_RArmF and LIG3Mut_RArm3R primers.

For the screening of the correctly targeted clones, LIG3pNeDaKI-EF1 x PGK-R2 and ZeoF1 x LIG3pNeDaKI-ER3 primer sets were used to ascertain the integrity of the left and right homology arms, respectively. After confirmation of a *KU86$^{flox/-}$*:*LIG3$^{mtio:NEO/flox}$* cell line, clones were transiently treated with Cre recombinase and subcloned to identify for *KU86$^{flox/-}$:LIG3$^{mito/-}$* clones. Screening of the Cre recombinase-treated clones was performed with LArmF x LIG3pNeDaKI-ER primers.

Construction of a *KU86^{floxl-}*:*LIG4$^{-/-}$* cell line

LIG4 targeting vectors were used as described (Oh et al. 2013) in two rounds of gene targeting with *KU86$^{flox/-}$* cells (Wang et al. 2009) to generate a *KU86*:*LIG4* doubly-mutant cell line.

LIG3 complementation

To construct the mitochondrial-only *LIG3* cDNA, the second and third ATGs in the ORF (open reading frame) of the *LIG3* cDNA were mutated to ATC. This mitochondrial-only *LIG3* cDNA was cloned into the pcDNA3.1(+) vector with a C-terminal HA epitope tag. For the nuclear-only *LIG3*, the N-terminal ORF that encodes the MLS was deleted and a FLAG-epitope tag was added to the C-terminus. This modified, nuclear-only *LIG3* cDNA was also cloned into a modified pcDNA3.1(+) vector, where the NEO gene had been replaced with a puromycin-resistance gene. Complementation constructs were linearized by PvuI restriction enzyme digestion and stably transfected into the relevant cell lines with Lipofectamine 2000. For selection, 1 mg/ml G418 and 2 μg/ml puromycin, respectively, were used.

Use of an inducible Cre system

A PiggyBac transposon system (Doherty et al. 2012) was used with slight modification. Cells were subcultured into 24-well plates a day before transfection. A vector expressing PiggyBac-transposase (0.4 μg) and a vector containing PiggyBac-CreERt2-transposon lacking a GFP marker (0.4 μg) were transfected using Lipofectamine 2000 according to the manufacturer's protocol. The transfected cells were subcultured 48 hr after transfection into a 96-well plate with 2 μg/ml of puromycin in the media for selection. Clones stably expressing CreERt2 were subsequently identified by immunoblotting using an antibody directed against Cre recombinase (data not shown). For CreERt2 induction, 10 nM of 4-hydroxytamoxifen (4-OHT), which was dissolved in ethanol, was used.

Immunoblotting

Whole cell extracts were prepared with RIPA buffer and 30 μg of protein was electrophoresed on a 4% to 20% gradient SDS (sodium dodecyl sulfate)-polyacrylamide gel and a rabbit, anti-human LIG3 monoclonal antibody (Gene Tex) was used at a 1:1000 dilution. HA (Covance) and FLAG (Sigma) antibodies were also used at a 1:1000 dilution. The actin antibody (Santa Cruz), which was used for the loading control, was diluted 1:250.

Immunocytochemistry (ICC)

Cells were plated on multi-chamber slides a day before analysis and subsequently fixed with 4% (v/v) paraformaldehyde in phosphate-buffered saline for 10 min. Slides were incubated in antigen retrieval buffer (100 mM Tris, 5% (v/v) urea, pH 9.5) at 95°C for 10 min. Permeabilization was performed with 0.1% Triton X-100. The LIG3 antibody was used at a 1:1000 dilution and an Alexa Fluor 488 goat, anti-mouse IgG antibody (Invitrogen) was used to visualize LIG3. DAPI (0.2 μg/ml) was used to stain the nucleus.

Etoposide and MMS (methyl methane sulfonate) sensitivity

An etoposide sensitivity assay was performed as described (Oh et al. 2013) with slight modifications. The cells were plated on a 6-well cell culture plate approximately 17 to 19 hr prior to drug treatment. Etoposide was dissolved in dimethyl sulfoxide to give a 10 mM stock solution. The cells were then incubated in etoposide-containing medium for 7 to 10 days, fixed and stained with crystal violet. For the MMS sensitivity test, cells were incubated in MMS-containing media for 1 hr and then maintained in drug-free media for 7 to 10 days. In all survival experiments the wild-type cells were grown for 7 days and the

LIG3 mutant lines were grown for 10 days to accommodate the latter's slower growth phenotype.

DNA EJing assays and plasmid rescue

The in vivo EJing reporter plasmid pEGFP-Pem1-Ad2 (Fattah et al. 2010) and the total A-NHEJ reporter pEJ2 (Bennardo et al. 2008) were used as described. HindIII- or I-SceI-digested pEGFP-Pem1-Ad2 or I-SceI-digested pEJ2 plasmid was co-transfected with a pCherry plasmid — used as a control of transfection efficiency — into the relevant cell line using Lipofectamine 2000. Green (EGFP) and red (Cherry) fluorescence were measured by FACS (fluorescence-activated cell sorting) 24 or 48 hr later. Repair efficiency was presented as a ratio of cells that were doubly positive for red and green over the number of cells that were only positive for red fluorescence.

For plasmid rescue experiments (Fattah et al. 2010), HindIII- or I-SceI-digested pEGFP-Pem1-Ad2 plasmid was transfected into the desired cell lines without the pCherry plasmid. Plasmids were rescued 24 hr after transfection using a mini-preparation protocol, transformed into E. coli and repaired plasmids were selected on LB plates containing 30 μg/ml of kanamycin. All the repair products were analyzed by sequencing, using a variety of primers located upstream and downstream of the Ad2 exon sequence.

Microhomology assay

A DNA microhomology assay was performed as described (Fattah et al. 2010). Cells were subcultured into 6-well plates a day before transfection. pDVG94 plasmid (2.5 μg) digested with EcoRV and AfeI was transfected using Lipofectamine 2000. Plasmid DNA

was recovered using a modified mini-preparation protocol at 24 to 48 hr after transfection. Repaired DNA junctions were PCR amplified using FM30 and 5'-radiolabeled DAR5 primers. PCR products were then digested with BstXI. Digested PCR products were separated by electrophoresis on a 6% polyacrylamide gel. The gel was then either dried and exposed to film, or stained with SybrGold and imaged on a Storm 840 fluorescent gel imager.

rAAV-mediated gene targeting

pAAV-HPRT-Puro, pAAV-helper, and pAAV-RC vectors were transfected into 70% confluent AAV-293 cells in 10 cm cell culture dishes and infectious rAAV-HPRT-Puro virus was harvested 3 days later by freeze/thawing as described (Kohli et al. 2004, Khan et al. 2011). The virus was subsequently purified using a rAAV virus purification kit and the viral titer was quantitated by qPCR. A day before infection, cells were subcultured into 2×10^5 cells/well in 6-well cell culture dishes in duplicate. Before adding virus, the exact number of cells was determined by counting one of the duplicate wells, and then virus at a MOI of 1×10^4 was added to the other well. Two days after infection, 1% of the cells were plated into a 10 cm culture dish without any drug selection and used to determine the plating efficiency. The remaining cells were plated into 10 cm culture dishes with 2 μg/ml puromycin. This medium was replaced 5 days later with puromycin-containing medium supplemented with 5 μg/ml of 6TG (6-thioguanine) except for one plate, which was used to quantitate the random integration frequency (i.e., those clones that were just G418 positive). When the cells had formed visible colonies at approximately 10 to 14 days later, the plates were fixed and stained with crystal violet.

Results

Generation of a conditionally-null *LIG3* HCT116 cell line

To generate a HCT116 cell line that was conditionally-null for *LIG3* expression, a rAAV gene-targeting methodology was adopted (Kohli et al. 2004, Khan et al. 2011), similar to the targeting strategy that was used to generate a *KU86* conditionally-null HCT116 cell line (Wang et al. 2009). The conditional targeting vector contained three LoxP sites flanking a *NEO* (neomycin-resistance) gene and exon 5 of *LIG3*, respectively (Fig. 2.1A). Eight correctly targeted first round clones ($LIG3^{NEO/+}$; Fig. 2.1B) from 210 G418-resistant clones were identified (relative gene targeting frequency: 3.8%). One of these clones was transiently treated with pGK-Cre to remove the NEO selection cassette (Fig. 2.1C). The resulting $LIG3^{flox/+}$ cell line was subjected to a second round of gene targeting using a non-conditional knockout vector, in which exon 5 was designed to be simply replaced by a floxed *NEO* gene (Fig. 2.1C). Ten correctly targeted second round clones were identified from 711 G418-resistant clones (relative gene targeting frequency: 1.4%). Additional analysis (data not shown) demonstrated that 9 of the clones were re-targeted (and therefore biologically uninteresting) whereas one clone was correctly targeted to the second allele (*i.e.*, $LIG3^{flox/NEO}$; Fig. 2.1D), which was subsequently infected with an AdCre virus to remove the *NEO* gene. The resulting $LIG3^{flox/-}$ cell line (Fig. 2.1E) was viable and when needed, Cre recombinase could be re-introduced to generate $LIG3^{-/-}$ cells (Fig. 2.1F).

PCR analyses were used to molecularly confirm the genetic designation of the cell lines. A 662 bp PCR product is diagnostic for the presence of the wild-type *LIG3* exon 5 whereas a 748 bp product should be generated when the floxed exon 5 DNA is used as a

substrate (Fig. 2.2A). As expected, PCR of the parental wild-type *LIG3*$^{+/+}$ cell line generated only the 662 bp product, whereas PCR of the *LIG3*$^{flox/+}$ cell line produced both the wild type (662 bp) and the floxed allele (748 bp) fragments (Fig. 2.2A). In contrast, LIG3 conditionally null cell lines — with or without the *NEO* selection cassette — generated only the 748 bp PCR product corresponding to the floxed allele (Fig. 2.2A).

To assess the conditionality of the cell line, *LIG3*$^{+/+}$ and *LIG3*$^{flox/-}$ cells were infected with an increasing amount of AdCre virus, and 5 days later genomic DNA was purified and analyzed by PCR and whole cell extracts were prepared and analyzed by immunoblotting. As the amount of Cre recombinase increased, the PCR signal from the genomic floxed allele (748 bp) decreased until the signal was undetectable (Fig. 2.2B). Similarly, at the protein level, the increasing presence of Cre completely ablated detectable LIG3 protein expression (Fig. 2.2C). In contrast, Cre expression in the parental wild-type *LIG3*$^{+/+}$ cells had no effect on formation of the exon 5-derived PCR product (662 bp) nor little effect on LIG3 protein expression (Fig. 2.2B and C). From these experiments, we concluded that we had successfully constructed a human *LIG3*$^{flox/-}$ (*i.e.,* conditionally null) cell line.

The mitochondrial form of *LIG3* is essential for human somatic cell survival

To assess the essential nature of *LIG3*, we initially infected the *LIG3*$^{flox/NEO}$ cell line (Fig. 2.1D) with AdCre and then isolated 22 individual G418-sensitive colonies by limiting dilution. Theoretically, two possible cell lines could have been recovered: *LIG3*$^{flox/-}$ and *LIG3*$^{-/-}$ (Fig. 2.1E and F, respectively), however all 22 recovered clones had a *LIG3*$^{flox/-}$ genotype and none were *LIG3*$^{-/-}$. This extreme asymmetry in the recovery of Cre-treated survivors strongly suggested that the *LIG3*$^{-/-}$ cell line was not viable.

Because the mitochondrial form of LIG3 is essential in mice (Gao et al. 2011, Ruhanen *et al.* 2011, Simsek et al. 2011), we tested whether this activity was conserved in human LIG3. To this end, we complemented the *LIG3*$^{flox/-}$ cells with a modified *LIG3* cDNA that could be expressed only in the mitochondria (mL3), and generated a stable *LIG3*$^{flox/-:mL3}$ cell line. The mL3 expression construct was made by mutating the second and third *LIG3* ATGs to ATCs (Fig. 2.3A). The nuclear form of LIG3 is normally translated from the second ATG, and without the second ATG only the longer mitochondrial-specific version of the protein should be made (Lakshmipathy and Campbell 1999). We mutated the third ATG simply as a precaution to ensure that no N-terminally truncated nuclear protein could be expressed. Importantly, after infecting LIG3$^{flox/-:mL3}$ cells with AdCre virus and isolating single cell clones by limiting dilution, we observed that 25 out of 43 were genotypically *LIG3*-null (*LIG3*$^{-/-:mL3}$; data not shown) — a result that contrasted sharply with 0 out of 22 obtained when we tried to establish *LIG3*-null cells in the absence of mL3 expression.

We next investigated whether the mitochondrial form of LIG3 could directly rescue the lethality of *LIG3*-deficient cells. When AdCre virus was used for this experiment, it was often difficult to differentiate the viral toxicity caused by the adenovirus from the apoptosis induced by the absence of LIG3. To improve upon this experimental set-up, a tamoxifen-inducible Cre recombinase was stably introduced into *LIG3*$^{flox/+}$, *LIG3*$^{flox/-}$, *LIG3*$^{flox/-:mL3}$ and LIG3$^{flox/-:nL3}$ (see below) cell lines. 4-OHT (4-hydroxyltamoxifen) treatment completely killed *LIG3*$^{flox/-}$ cells (clone #5) but had little effect on *LIG3*$^{flox/+}$ cells (clones #101-1 and #101-21), which still had a functional LIG3 allele even after Cre treatment. Significantly, the lethality of *LIG3*-null cells was significantly rescued by expression of the mitochondrial form of *LIG3* (Fig. 2.3B; clones #73-5 and #73-6). From these experiments we concluded that the mitochondrial form of LIG3 is essential for human somatic cell viability.

Finally, a mitochondrial-exclusive expression pattern of mL3 was verified by immunohistochemistry. In the parental wild-type HCT116 *LIG3*$^{+/+}$ cells, LIG3 protein was expressed ubiquitously throughout the cell, whereas in *LIG3*$^{-/-:mL3}$ cells fluorescent signal was detected virtually exclusively in the cytoplasm (Fig. 2.3C).

Complementation of *LIG3*$^{-/-:mL3}$ with a nuclear LIG3 cDNA

With a viable *LIG3*-null cell line in hand (*i.e.*, *LIG3*$^{-/-:mL3}$ cells), we were equipped to investigate the phenotypes resulting from the loss of nuclear LIG3 expression. Before beginning these analyses, however, we augmented our reagents with a derivative cell line that re-expressed a nuclear-specific *LIG3* cDNA. For this approach, a nuclear-only *LIG3* cDNA, nL3, was generated by deleting the N-terminal MLS from the wild-type cDNA and by adding a C-terminal FLAG epitope-tag (Fig. 2.3A). We isolated stable *LIG3*$^{-/-:nL3}$ and *LIG3*$^{-/-:mL3:nL3}$ cell lines, the later of which exhibited a strong LIG3 ICC signal from the nucleus in addition to pan-cytoplasmic staining (Fig. 2.3C). Importantly, the nuclear-exclusive form of LIG3 — in contrast to the mitochondrial-exclusive form — was incapable of rescuing the lethality of *LIG3*-nulls (Fig. 2.3B; clone #12-12).

A *LIG3* deficiency causes a growth defect

While the mitochondrial form of LIG3 rescued the lethality of *LIG3*-deficient cells, the rescued phenotype was not as robust as even *LIG3* heterozygous cells (Fig. 2.3B, compare clones #73-5 and #73-6 with #101-1 and #101-21). Thus, we investigated whether there was a growth defect associated with the absence of LIG3. Three thousand cells were seeded on day 0 into each well of a 6-well tissue culture dish and the number of cells in each well was counted on days 4 to 8 (Fig. 2.4A). LIG3$^{flox/-}$ cells showed a slight

haploinsufficiency for growth, which was significantly exacerbated in the LIG3$^{-/-:mL3}$ cell line. From these experiments we concluded either that the absence of nuclear LIG3 — or inadequate mitochondrial LIG3 — expression results in a proliferation defect.

LIG3-null cells are not sensitive to DNA damaging agents

Given that *LIG3* has been implicated in both SSB (single-strand break) and DSB repair pathways (Wang et al. 2005, Caldecott 2007), colony-forming assays were used to examine the sensitivity of LIG3-null cells to a variety of DNA damaging agents. In these survival experiments, the wild-type cells were grown for 7 days and the *LIG3* mutant lines were grown for 10 days to accommodate the latter's slower growth phenotype. Etoposide is a topoisomerase II inhibitor and a powerful radiomimetic drug that induces DNA DSBs (Pacchierotti and Ranaldi 2006). As a positive control, a *LIG4*$^{-/-}$ cell line (Oh et al. 2013) was, as expected, exquisitely sensitive to etoposide, even at the lowest concentration (Fig. 2.4B). In contrast, *LIG3*$^{flox/-}$, *LIG3*$^{-/-:mL3}$ and *LIG3*$^{-/-:mL3:nL3}$ did not exhibit any increased sensitivity to etoposide compared to the wild-type parental cells (Fig. 2.4B). Similarly, the LIG3-deficient cell lines were slightly, but not significantly, sensitive to MMS (methyl methanesulfonate), an alkylating agent that induces SSBs at low doses and DSBs at high doses (Fig. 2.4C). In conclusion, the absence of LIG3 (and presumably the A-NHEJ DSB repair pathway) did result in a slower growth phenotype, but this was not reflected in a hypersensitivity to DNA damaging agents.

The absence of LIG3 does not affect the overall DNA EJing activity of human cells

The total DNA EJing activity of LIG3-null cells was measured using pEGFP-Pem1-Ad2, an extra-chromosomal reporter assay vector (Seluanov *et al.* 2004, Wang et al. 2006,

Fattah et al. 2010). The reporter plasmid consists of the GFP gene, which is interrupted by a 2.4 kb intron derived from the rat Pem1 gene. An exon (Ad2) derived from adenovirus serotype 2 has been introduced into the middle of the intron and is flanked by *HindIII* and *I-SceI* restriction enzyme recognition sites (Fig. 2.5A). Without modification, the GFP gene is not expressed because the Ad2 exon is incorporated into GFP mRNA (Fig. 2.5C). Pre-digestion of the plasmid with *HindIII* or *I-SceI* removes the Ad2 exon and generates a linear plasmid with compatible (*i.e.*, ends that can be joined by simple ligation) or incompatible (*i.e.*, ends that require some sort of processing before they are rejoined) ends, respectively (Fig. 2.5B). Productive end joining of the linear plasmid after it is transfected into the experimental cell line can be quantitated using FACS analysis of GFP expression (Fig. 2.5C).

When the pEGFP-Pem1-Ad2 plasmid was transfected into the parental wild-type cell line, approximately 35% of the plasmid was productively end joined regardless of whether it had been digested by either *HindIII* or *I-SceI* (Fig. 2.5D). As expected, *LIG4*-null cells were profoundly impaired in this DNA EJing activity and showed only a few percent of activity. In contrast, $LIG3^{flox/-}$ and $LIG3^{-/-:mL3}$ cell lines had DNA EJing activities indistinguishable from wild-type cells (Fig. 2.5D). Thus, partial or complete deficiencies in nuclear *LIG3* did not appear to affect the overall DNA EJing activity of human somatic cells.

Microhomology-mediated EJing is still detectable in human somatic cells lacking nuclear LIG3 expression

Even though the overall DNA EJing activity was unchanged in *LIG3*-null cells, it was still possible that distinct DNA repair pathways were being used. To examine this

possibility, we next quantitated microhomology-mediated EJing. To this end, a reporter substrate, pDVG94 that is biased towards detecting microhomology-mediated EJing events (Verkaik *et al.* 2002, Weterings *et al.* 2009, Fattah et al. 2010, Oh et al. 2013) was used *in vivo* to measure this activity of LIG3-deficient cells. *EcoRV* and *AfeI* digestion of pDVG94 generates a blunt-ended, linear, double-stranded substrate with 6 bp direct repeats at both ends (Fig. 2.6A). Repair of this substrate by C-NHEJ generally generates a product that retains at least some of either repeat, whereas microhomology-mediated EJing (*i.e.*, A-NHEJ) produces a unique product that contains only a single repeat, and which now forms the recognition sequence for the *BstXI* restriction enzyme. The linearized pDVG94 plasmid was introduced into the relevant cell lines and 24 hr later the DNA was recovered from the cells and repaired junctions were amplified by PCR with radiolabeled primers (Fig. 2.6A). The resulting ~180 bp PCR products were then digested with *BstXI* and subjected to agarose gel electrophoresis. *BstXI*-resistant DNA corresponds to C-NHEJ-mediated repair events whereas a 120 bp product is diagnostic of A-NHEJ/microhomology-mediated EJing. Wild-type cells generated only ~1% of the 120 bp product (Fig. 2.6B), consistent with most of the EJing in human somatic cells resulting from C-NHEJ (Fattah et al. 2010, Oh et al. 2013). Similarly, and again as expected (Fattah et al. 2010, Oh et al. 2013), a LIG4-null cell line showed highly elevated levels of the 120 bp product indicative of a virtually exclusive reliance on A-NHEJ (Fig. 2.6B). In contrast, in either the $LIG3^{flox/-}$ or two independent $LIG3^{-/-:mL3}$ cell lines (clones #29 and #35) the amount of the 120 bp product was either slightly elevated or unchanged in comparison to wild-type cells (Fig. 2.6B), indicating that *LIG3*-deficient cells were still able to carry out microhomology-mediated EJing.

A LIG3 deficiency does not affect the overall rAAV-mediated gene-targeting rate

AAV infections in humans are apathogenic and this feature has made rAAV-mediated gene targeting technology one of the more promising candidates for therapeutic use (Khan et al. 2011). The utility of rAAV, however, is offset by random viral integration events, which are not desired and potentially mutagenic. Thus, increasing the overall correct gene targeting frequency is one of the most sought-after advances for rAAV-mediated gene targeting technology and for gene therapy in general. Previous studies from our laboratory have indicated that neither HR nor C-NHEJ was responsible for rAAV random integration events (Fattah et al. 2008, Oh et al. 2013). In addition, sequencing results from other laboratories had indicated that rAAV random integration events could be mediated by microhomology usage (Miller *et al.* 2005, Nakai *et al.* 2005). Together, these observations led to the hypothesis that rAAV random integrations are mediated by A-NHEJ, and to the prediction that disruption of A-NHEJ (*e.g.*, by functionally inactivating LIG3) would ablate random integrations and thus improve the correct rAAV-mediated gene-targeting rate.

To test this hypothesis, we performed gene targeting in a *LIG3*-null cell line at the HPRT (hypoxanthine phosphoribosyl transferase) locus using a rAAV gene targeting vector designed to disrupt exon 3 of *HPRT*. *HPRT* is an X-chromosome linked gene and the enzyme encoded by *HPRT* is needed to generate nucleotides through the purine salvage pathway (Hladnik *et al.* 2008). Cells expressing a wild-type *HPRT* gene are lethally poisoned by the toxic nucleoside analog 6-TG (6-thioguanine), whereas cells with a defective *HPRT* gene can survive in the presence of 6-TG. Because HCT116 was derived from a male patient, it contains only a single X-chromosome and therefore after a single round of gene targeting, 6-TG selection could be used to isolate correctly targeted clones. Interestingly, *LIG3*-null cells showed a 2-fold increase in the frequency of correct

targeting (Fig. 2.7B). Unexpectedly, however, the frequency of random integration events was not reduced in *LIG3*-null cells, but was also actually enhanced (Fig. 2.7A). Consequently, there was no statistical difference in the relative gene-targeting rate between the wild type and *LIG3*-null cell lines (Fig. 2.7C). These data demonstrated (quite unexpectedly) that while *LIG3* is a general suppressor of rAAV integrations, it does not preferentially affect random versus correct targeting events.

The absence of Ku reveals a requirement for *LIG3*

The above experiments demonstrated that while the absence of *LIG3* resulted in a slower growth phenotype, it did not *i*) make the cells hypersensitive to DNA damaging agents, *ii*) manifest itself in any detectable DNA DSB repair deficiency nor *iii*) deleteriously impact on the process of gene targeting. Our laboratory (Fattah et al. 2008) and others (Guirouilh-Barbat *et al.* 2007, Bennardo et al. 2008, Mansour *et al.* 2008, Schulte-Uentrop *et al.* 2008) have shown that cells reduced or deficient for the Ku heterodimer show a marked increase in the activity of A-NHEJ. We therefore reasoned that if LIG3 is mediating A-NHEJ repair reactions that are too rare to detect in an otherwise wild-type cell, then there might be an observable effect of the absence of nuclear LIG3 in a *KU86*-null cell. To test this hypothesis, we constructed a $KU86^{flox/-}$:$LIG3^{mito/-}$ doubly mutant human cell line. This was achieved by "knocking-in" ATG > ATC mutations into one endogenous *LIG3* allele, and knocking out the second *LIG3* allele by rAAV-mediated gene targeting in a conditionally-null *KU86* cell line (Fig. 2.8A). After also engineering in the above-mentioned tamoxifen-inducible Cre recombination system, we confirmed by immunofluorescence (Fig. 2.8B) and PCR (Fig. 2.8C) that the $KU86^{flox/-}$:$LIG3^{mito/-}$ cells only expressed mitochondrial *LIG3* and were conditionally-null for *KU86*, respectively.

Furthermore, in order to directly compare *LIG3* to another DNA ligase in this context, we also created a *KU86*$^{flox/-}$:*LIG4*$^{-/-}$ Cre-inducible cell line (Fig. 2.8C). Although both doubly mutant cell lines eventually die due to Ku-regulated telomere defects (Wang et al. 2009), they do survive 3 to 7 days post Cre treatment, which was a sufficient time window to assay the levels of DNA repair in the cells.

Consequently, we performed a microhomology assay using the aforementioned pDVG94 reporter plasmid (Fig. 2.9A). We confirmed our previously published finding (Fattah et al. 2010) that in the absence of the Ku, there is a large (~20-fold) increase in the relative amount of A-NHEJ repair (Fig. 2.9B). Importantly, the absence of nuclear *LIG3* in the *KU86*-deficient cells completely abrogated this increased A-NHEJ (compare *KU86*$^{-/-}$ with *KU86*$^{-/-}$:*LIG3*$^{mito/-}$; Fig. 2.9B). Notably, although the absence of *LIG4* by itself resulted in exceedingly few total repair events (Fig. 2.9D), those that did occur were almost exclusively biased towards microhomology and the presence or absence of Ku in the cells had no discernible effect on this bias (compare *LIG4*$^{-/-}$ with *KU86*$^{-/-}$:*LIG4*$^{-/-}$; Fig. 2.9B). Interestingly, the converse was not true. Thus, the absence of LIG4 in the cells always shifted the repair bias towards A-NHEJ regardless of the status of Ku expression (compare *KU86*$^{-/-}$ with *KU86*$^{-/-}$:*LIG4*$^{-/-}$; Fig. 2.9B).

The pDVG94 microhomology assay measures relative A-NHEJ activity. To determine the impact of ligation on absolute levels of A-NHEJ, we performed an additional analysis with the aforementioned cell lines using the pEJ2 reporter plasmid. This plasmid contains an *I-SceI* restriction enzyme site flanked by 8 bp of microhomology situated between a promoter and GFP coding sequences (Fig. 2.10A; (Bennardo et al. 2008)). In addition, stop codons have been engineered in all three reading frames to disrupt the translation that starts at an N-terminal epitope tag. Thus, this plasmid can be linearized

by *I-SceI* restriction enzyme digestion and transfected into cells. Microhomology-mediated repair removes the stop codons and allows GFP expression, whereas repair by C-NHEJ does not. Consequently, cells were transfected with *I-SceI*-digested pEJ2 plasmid and mCherry (as a transfection control) and assayed for A-NHEJ activity. First, it was obvious that in all the cell lines that A-NHEJ is a rare event (Fig. 2.10B), consistent with our other analyses. Nonetheless, the absence of Ku caused a 4-fold increase in GFP-positive cells and almost all of the increase could be abrogated by the absence of LIG3, but not LIG4 (Fig. 2.10C).

In toto, these data clearly demonstrated a separation of function between the two ligases and strongly suggested that nuclear LIG3 does play a significant role in A-NHEJ, but only in the absence of Ku.

In the absence of Ku, most EJing remains *LIG4*-dependent

All of the above data were consistent with an old model in which the vast majority of EJing events in human somatic cells are Ku- and LIG4-dependent and that only a minority of repair events are normally repaired through a Ku-suppressible, LIG3-dependent pathway (Fig. 2.11A). This model predicts that in a Ku-deficient human cell, the total EJing events should be sensitive to the status of LIG3 and unaffected by the presence in LIG4. To test this prediction, we quantitated the total amount of EJing in our cell lines using the pEGFP-Pem1-Ad2 reporter plasmid (Fig. 2.12A). The *KU86*$^{-/-}$, *LIG3*$^{-/-:mL3}$ and *KU86*$^{-/-}$:LIG3$^{mito/-}$ cell lines had virtually indistinguishable robust total EJing activity whereas a Ku86$^{-/-}$:LIG4$^{-/-}$ cell line was extremely deficient (Fig. 2.12B). Thus, very unexpectedly, in a Ku-deficient human cell line, the vast majority of EJing events remain *LIG4*-dependent.

To extend these findings, we recovered the repaired pEGFP-Pem1-Ad2 plasmids out of cells 24 hr post transfection. These plasmids were analyzed by quantifying the plasmids that had simply religated their HindIII-linearized ends ("perfect rejoins"; (Fattah et al. 2010)). The repaired plasmids recovered from cells were amplified by colony PCR, and the products were then re-digested with HindIII and analyzed by gel electrophoresis (Fig. 2.12C). As we had previously observed (Fattah et al. 2010), the absence of Ku increased the frequency of perfect rejoins (Fig. 2.12D), however the absence of LIG3 had no impact on this process (compare *KU86*$^{-/-}$ with *KU86*$^{-/-}$:*LIG3*$^{mito/-}$; Fig. 2.12D). Interestingly, *KU86*$^{-/-}$:*LIG4*$^{-/-}$ (and *LIG4*$^{-/-}$) cells carried out almost exclusively perfect rejoining, implicating therefore, by default, LIG1 in this process (Paul *et al.* 2013).

Altogether, these data demonstrated that *LIG4* is still required for most of the Ku-independent EJing events in human cells. Moreover, they demonstrate that *LIG3* is required for much of the Ku-suppressible A-NHEJ activity. And finally, the data suggest that at least some of the perfect rejoining and A-NHEJ observed in either wild type or Ku-deficient cells can be mediated by an additional pathway, which we infer perforce to be *LIG1*-mediated (Paul et al. 2013).

Discussion

LIG3 has an essential mitochondrial function

We have generated a viable human somatic cell line that lacks the expression of nuclear LIG3. Somewhat paradoxically, we nonetheless demonstrate that *LIG3* is an essential gene. Thus, no living null cells could be isolated after an AdCre infection of the *LIG3*$^{flox/-}$ cell line. In contrast, after complementing the *LIG3*$^{flox/-}$ cell line with a mitochondrial-specific cDNA isoform of LIG3, we readily isolated viable *LIG3*-null (*LIG3*$^{-/-:mL3}$) cells. Therefore, *LIG3* is dispensable in the nucleus of human cells but essential in mitochondria.

A similar conclusion that *LIG3* is essential for mitochondrial function was reached using the chicken cell line, DT40 (Arakawa et al. 2012), and in the mouse (Puebla-Osorio et al. 2006, Gao et al. 2011, Simsek et al. 2011). What the essential activity of mitochondrial LIG3 is, however, is unclear. One obvious function would be a requirement for LIG3 in Okazaki fragment maturation during mitochondrial DNA replication. LIG1 mediates Okazaki fragment maturation in the nucleus (Ellenberger and Tomkinson 2008, Zheng and Shen 2011), but since LIG1 lacks a MLS and it is not detected in mitochondria, it is clear that some other ligase must perform this function in mitochondria (Arakawa et al. 2012, Le Chalony *et al.* 2012). Similarly, since LIG4 also lacks a MLS, is non-essential, and appears to be involved exclusively in C-NHEJ, it is also a poor candidate (Oh et al. 2013). In contrast, the mitochondrial-specific isoform of LIG3 should be able to mediate the ligation of Okazkai fragments, which are very similar to the repair intermediates that LIG3 is known to ligate together during BER (Caldecott et al. 1994, Frosina *et al.* 1996).

rAAV random integrations are not mediated solely by LIG3/A-NHEJ

Correct gene targeting events require HR, and the (much more frequent) random integrations of the targeting vector are presumably facilitated by DNA EJing pathways (*i.e.*, C-NHEJ, A-NHEJ or both). Indeed, large-scale whole-genome sequencing results demonstrated that rAAV random integration events were mediated by some form of NHEJ (Miller et al. 2005, Nakai et al. 2005). Reductions, however, in the expression of Ku did not significantly impact the frequency of random integrations of rAAV vectors during gene targeting (Fattah et al. 2008). Similar observations have been made for *LIG4*-null cells (Oh et al. 2013) suggesting strongly that C-NHEJ is not required for this process. Together, these observations led to the hypothesis that rAAV random integrations are instead mediated by A-NHEJ, and to the prediction that disruption of A-NHEJ (*e.g.*, by functionally inactivating *LIG3*) would significantly improve the correct rAAV-mediated gene-targeting rate.

Surprisingly, no such effect was observed (Fig. 2.7) and the random rAAV integration frequency was actually (~3-fold) elevated, demonstrating that *LIG3* could even be viewed as a suppressor of these events. A parsimonious explanation for these results is that if random rAAV integrations are neither mediated by *LIG4* nor *LIG3* then they may be mediated by *LIG1* (see also Fig. 2.11B). Although LIG1 is predominately thought of as a DNA replication-linked ligase, it is enzymatically capable of resealing the staggered ends of the DSBs generated during random rAAV integration (Ellenberger and Tomkinson 2008). Unfortunately, LIG1 is an essential gene in the HCT116 cell line (unpublished data), which precludes gene targeting experiments and a direct test of the hypothesis.

Alternatively, we propose that in the absence of any one particular ligase that rAAV integration is performed by the remaining ligases. There is precedent from other cellular

activities for the mammalian ligases being redundant. This is especially true of LIG1 and LIG3. Thus, viable murine *LIG1*-null MEFs actually proliferate rather normally, suggesting that some other ligase must be compensating for LIG1 during DNA replication (Bentley *et al.* 1996, Bentley *et al.* 2002, Le Chalony et al. 2012). Similarly, the co-depletion of *LIG1* in murine *LIG3*-knockout cells sensitized the cells to MMS exposure, an activity previously thought to be solely the purview of LIG3 (Gao et al. 2011). Perhaps the most compelling case for genetic redundancy comes from work in the chicken DT40 cell line, where single, double and or triple mutants were generated to establish that *LIG1* and *LIG3* are functionally redundant for viability (Arakawa et al. 2012). A similar conclusion that *LIG1* and *LIG3* are functionally redundant for *in vitro* end joining in human extracts was obtained by reducing the levels of each, or both, protein(s) using RNA interference or neutralizing antibodies (Liang *et al.* 2008). If this redundancy between *LIG1* and *LIG3* extends to rAAV random integration it provides a plausible explanation for the lack of a compelling phenotype for this process in the LIG3-null human cells.

LIG3 is required for A-NHEJ events normally suppressed by Ku

In order to generate a 120 bp *BstXI*-dependent restriction product from the pDVG94 plasmid a unique repair event is required: both ends of the blunt-ended plasmid must be resected and the 6 nt long complementary strands must be precisely annealed and ligated. This repair product can be detected only at low levels (a few percent of the total) in wild-type cells and has been widely interpreted as being produced by A-NHEJ (Verkaik et al. 2002, Fattah et al. 2010). Moreover, when cells are mutated for any of the canonical C-NHEJ genes, the 120 bp fragment becomes virtually the sole repair product (Fig. 2.6 and 2.9; (Fattah et al. 2010)). All of these observations led to the prediction that the ablation

of A-NHEJ should block the formation of this repair product. However, the appearance of the 120 bp product was completely unaffected by the absence of nuclear LIG3 (Fig. 2.6 and 2.9). A trivial explanation is that the cell line is leaky and expresses some LIG3 in the nucleus. While we cannot unequivocally rule out this possibility, we do not believe that this is a likely explanation. First, we mutated not only the normally-used ATG for the nuclear form of LIG3, but a downstream ATG as well to ensure that not even truncated forms of the protein would be expressed. In addition, a mitochondrial-exclusive expression pattern of the mL3 construct was confirmed by immunocytochemistry (Fig. 2.3C and 2.8B). Still if mL3 protein did leak into nucleus, even at undetectable levels, this would likely have functional relevance since even low levels of LIG3 are sufficient for effective A-NHEJ (Windhofer *et al.* 2007). Finally, we also note that all of our repair data was generated using episomal reporter vectors that may not always faithfully represent endogenous DNA repair events, which are, perforce, chromosomal in nature (Mao *et al.* 2008).

An alternative attractive possibility is that LIG3 is redundant with LIG1 in the A-NHEJ pathway (Katyal and McKinnon 2011). This hypothesis, which we elaborated above as a possible explanation for the absence of an anticipated effect on gene targeting, would explain the continued production of the 120 bp product even in the absence of LIG3. In almost every experimental situation where the expression of both *LIG1* and *LIG3* has been reduced, the resultant repair activity is significantly less than when either ligase is individually reduced (Liang et al. 2008, Gao et al. 2011, Arakawa et al. 2012). Thus, our data suggests that there are redundant LIG1- and LIG3-dependent forms of A-NHEJ, although only the LIG3-branch appears to be suppressible by Ku (see also Fig. 2.11).

It is unclear what biological function this Ku-suppressible, LIG3-dependent A-NHEJ activity serves. To our knowledge, the only nuclear DSB "activity" that can currently be ascribed to LIG3 in C-NHEJ-proficient cells is the ability to facilitate chromosomal translocations (Simsek et al. 2011); an activity that seems untenable as an evolutionary explanation for the presence of LIG3. We note, however, that although nuclear LIG3-null cells exhibit no obvious DSB repair phenotypes *per se*, they do exhibit a growth defect, implying that the absence of LIG3 is deleterious for the normal progression of cell division. However, this requirement for LIG3 could simply be related to its role in BER, an activity likely needed for efficacious cell cycle progression (Krokan and Bjoras 2013). In addition, we propose that there may be transient periods in the cell cycle when Ku's activity is suppressed, and where LIG3-dependent A-NHEJ can be engaged in a beneficial manner. An obvious time for this would be during S-phase. Thus, cells normally require HR to accurately re-start/repair stalled/broken replication forks (Ghosal and Chen 2013), and the repression of C-NHEJ via the suppression of Ku is likely instrumental in this process. The DNA ligase required for these HR transactions is unknown but our data is consistent with the recent suggestion that LIG3 may be required for at least Fanconi anemia pathway-mediated HR processes (Huang and Li 2013). Whether LIG3-dependent A-NHEJ activities could also be utilized in this situation is clearly a question that deserves additional investigation.

A new model for EJing in human somatic cells

Previous studies of EJing in human cells had postulated a model of DSB repair that consisted of a predominant Ku-dependent, LIG4-dependent C-NHEJ pathway and a minor Ku-suppressible, LIG3-dependent pathway (Fig. 2.11A; (Bennardo et al. 2008, Fattah et

al. 2010, Boboila et al. 2012)). Here we demonstrate that this model is, at best, over-simplified. Thus, our data show that there is a constant low level of A-NHEJ in human cells that is unaffected by the presence or absence of LIG3 (Fig. 2.6 and 2.10). We ascribe this activity to LIG1 and propose that the LIG1-dependent branch of A-NHEJ is separable from the LIG3-dependent branch (Fig. 2.11B). Moreover, we demonstrate that the LIG3-dependent branch of A-NHEJ is uniquely Ku-suppressible (Fig. 2.9 and 2.11B). Most interestingly, however, this work has demonstrated that there is a large fraction of EJing events that are Ku-independent, but nonetheless LIG4-dependent (Fig. 2.12). In a previous study we had demonstrated a robust EJing activity in Ku-deficient human somatic cells and had inferred that this activity was likely due to LIG3-dependent A-NHEJ (Fattah et al. 2010). However, the experiments presented here show that little of this Ku-independent activity is actually LIG3-dependent, but is instead LIG4-dependent. Interestingly, but in retrospect not unexpectedly, this Ku-independent, LIG4-dependent activity is capable of performing sticky-ended ligations at very high frequency (Fig. 2.12C; (Fattah et al. 2010)). Thus, we propose that there may a hierarchical order to C-NHEJ in which direct ligation of repair events is first attempted and only if that fails is Ku recruited to enable more complicated downstream processes such as end resection and polymerization before ligation ensues (Fig. 2.11B). Our postulated model is consistent with the recent demonstration that XRCC4 and XLF can facilitate the bridging of DNA ends for subsequent LIG4-dependent ligation (Mahaney *et al.* 2013).

In summary, these studies have revealed that human LIG3 is an essential gene due to its requirement in mitochondrial function. Moreover, we demonstrate that while nuclear LIG3 is not required for random gene targeting integrations, it is required for mediating the enhanced A-NHEJ repair events observed in the absence of Ku. Finally, we have

identified a novel EJing activity observed in the absence of Ku that is LIG4-dependent. Altogether, these observations enhance our understanding of EJing in human somatic cells.

Tables

Table 2.1. List of PCR primers

Name of the primers	Sequences
Exon5_LARM_F1	5'-ACATAAGCGGCCGCAGAGCACTTTGGCATCTGTCTTC-3'
Exon5_LARM_SacR1	5'-GGCGGCCCGCGGAAAAAATTAAAAAATTAGCTGG-3'
Exon5_RARM_XhoF1	5'-CGCTCGAGGGCTTTTATTCTGGACTCTTTTTTTC-3'
Exon5_RARM_R1	5'-ACATAAGCGGCCGCTGGAGTAGGCAAGAGACTCATAC-3'
Exon5_KpnF1	5'-CCGGTACCGTAGAGATGGGGTCTTTCTTTGTTGC-3'
Exon5_XhoR1	5'-CGCTCGAG***ATAACTTCGTATAATGTATGCTATACGAAGTTAT***CCAGGAGAG ACAGAGGGGGCAAG-3' (*Bold italic indicates the LoxP sequence)
Exon5_SC_F2	5'-ATGAGCATCCTGAATAGGCCTTTCCTCCGG-3'
NeoR2	5'-AAAGCGCCTCCCCTACCCGGTAGGGCG-3'
LIG3_LArm_F3	5'-TGCCACCATGTCCAGCTAA-3'
LIG3_RArm_R2	5'-GAGTCCAGAATAAAAGCC-3'
Lig3Mut_LArmF	5'-CCCTGCTTTAGAAATAAGAAAACAGATGCAG-3'
Lig3Mut_LArmR	5'-CATAAAAGCTCTTTTTAGGCCATCAGCAAC-3'
Lig3Mut_RArmF	5'-GGATTTTCTTCATTGTGATGTACTAGATGTACATGTGTGC-3'
Lig3Mut_RArm3R	5'-GCAGTTGAACTCCTTATCCCTTTTTCTCTC-3'
Lig3pNeDaKI-EF1	5'-CACAACCAGCGCATTCATGGATTC-3'
PGK-R2	5'-CCGGTGGATGTGGAATGTGT-3'
ZeoF1	5'-GGACGACGTGACCCTGTTCATCAG-3'
Lig3pNeDaKI-ER3	5'-CAGGATGACCTGAGTCCAGCAGTGC-3'
LArmF	5'-CCCTGCTTTAGAAATAAGAAAACAGATGCAG-3'
Lig3pNeDaKI-ER	5'-CGAATGCAACATTAAACACTGGCC-3'
FM30	5'-CTCCATTTTAGCTTCCTTAGCTCCTG-3'
DAR5	5'-TGCTTCCGGCTCGTATGTTGGTTGGAAT-3'

Figures

Figure 2.1

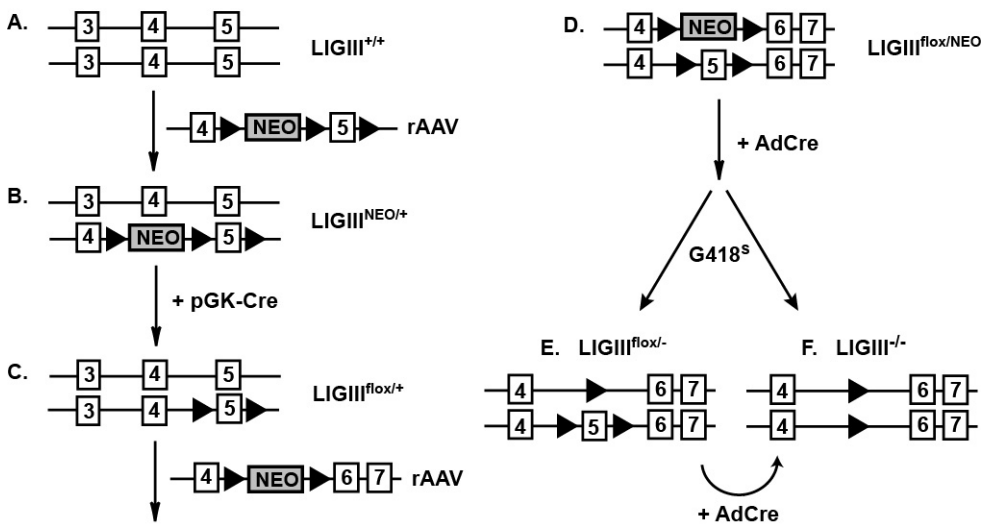


Figure 2.1. Scheme for the functional inactivation of the human *LIG3* locus.

A) A cartoon of part of the LIG3 genomic locus (*LIG3*$^{+/+}$). The numbered rectangles represent exons and lines in between are introns. The conditional knockout vector has three loxP sites (filled triangles) that flank a NEO cassette and exon 5, respectively. B) After the first round of targeting, G418-resistant clones were selected and among them, correctly targeted clones (*LIG3*$^{NEO/+}$) were identified by PCR. The *LIG3*$^{NEO/+}$ clones were treated with pGK-Cre to remove the NEO gene prior to the second round of targeting, which generated LIG3$^{flox/+}$clones. C) A G418-sensitive *LIG3*$^{flox/+}$ clone was targeted with an exon 5 knockout vector. D) A correctly targeted clone, LIG3$^{flox/NEO}$, was treated with AdCre to derive a G418-sensitive *LIG3*$^{flox/-}$ clone, depicted in E). AdCre treatment of *LIG3*$^{flox/NEO}$ can also generate a LIG3$^{-/-}$ clone, depicted in (F), if the recombination happens in both alleles. F) AdCre treatment of *LIG3*$^{flox/-}$ clones results in *LIG3*$^{-/-}$ cells.

Figure 2.2

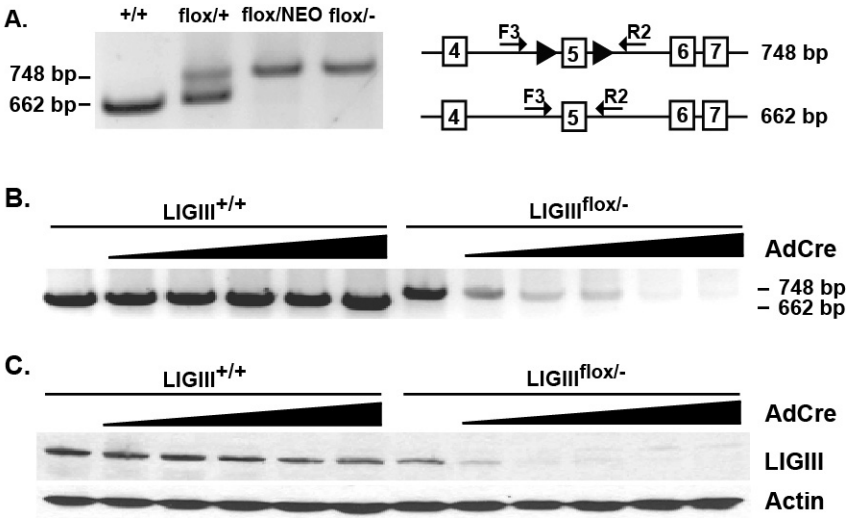


Figure 2.2. Generation of a *LIG3* conditionally-null cell line.

A) PCR confirmation of LIG3flox/− cells. A primer set, LIG3_LArm_F3 (F3) and LIG3_RArm_R2 (R2), which yield differently sized PCR products from the floxed allele (748 bp) and wild-type allele (662 bp) were used. With the same primer set, an allele containing the NEO gene generated a band over 3 kb and the loxP-only allele yielded a 97 bp product, but they are not shown in this figure. B) The floxed allele disappears after AdCre infection. With increasing amounts of AdCre virus, the 748 bp PCR product derived from the floxed allele disappeared in the LIG3flox/− clone. In contrast, the expression of Cre did not affect the LIG3+/+ control cells. C) Western blot analysis confirms the loss of LIG3 protein. Protein samples were produced in parallel to the DNA samples in (B). Increasing amount of AdCre virus correlated with decreasing amount of LIG3 protein only in the experimental cell line.

Figure 2.3

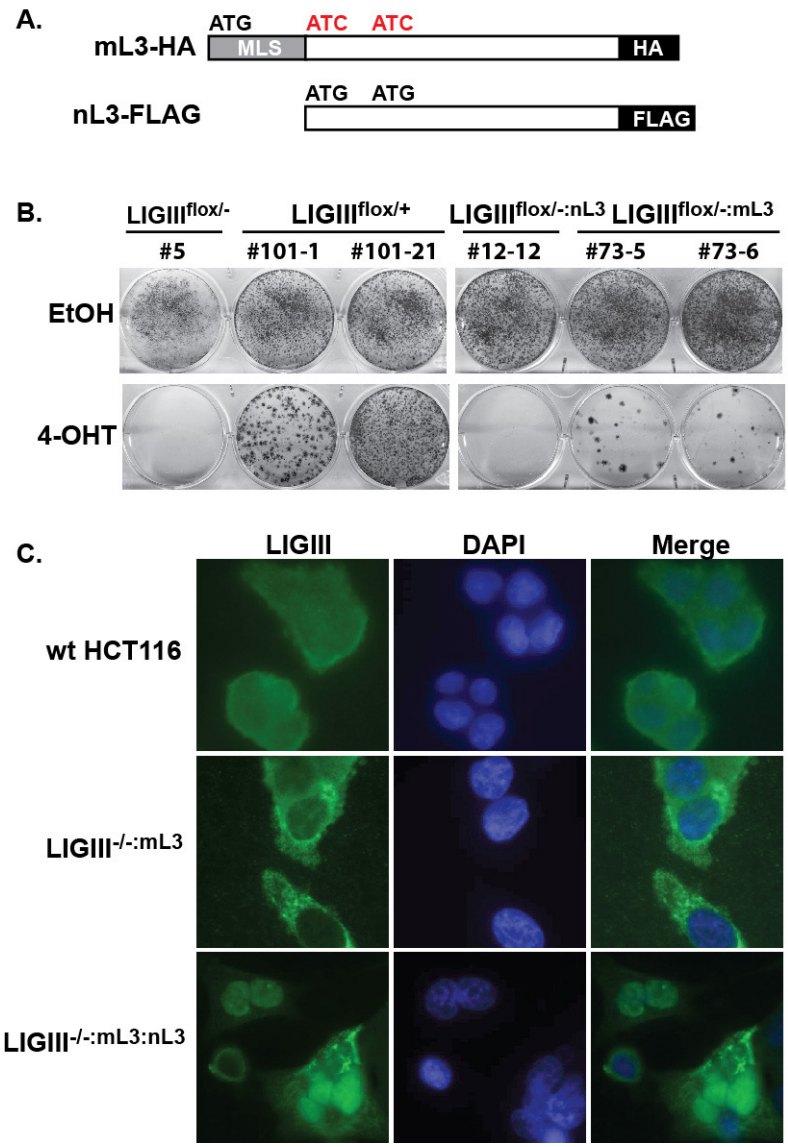


Figure 2.3. Complementation with a mitochondrial-only *LIG3* cDNA rescues the lethality of *LIG3*$^{-/-}$ cells.

A) Constructs used for complementation. To make a mitochondrial-only *LIG3* (mL3), the second and third ATGs were mutated to ATCs. For the nuclear-only *LIG3* (nL3) construction, the N-terminal MLS sequence was deleted. mL3 and nL3 were subsequently

modified with C-terminal HA- and FLAG- epitope tags, respectively. B) Cre-recombinase linked to the estrogen receptor (CreERt2) was stably expressed in *LIG3*$^{flox/+}$, *LIG3*$^{flox/-}$, *LIG3*$^{flox/-:mL3}$ and *LIG3*$^{flox/-:nL3}$ cell lines. #101-1 and #101-21 are two independent *LIG3*$^{flox/+:CreERt2}$ clones. #73-5 and #73-6 are two independent *LIG3*$^{flox/-:mL3:CreERt2}$ clones. Cells were plated onto 6-well plates and a day afterwards 4-OHT was added to induce Cre recombination. Cells were maintained either in 4-OHT containing media or EtOH-containing media, which served as a negative control, for 2 to 3 weeks until they formed colonies, which were then fixed and stained by crystal violet. C) Nuclear and mitochondrial localization of *LIG3* in complemented clones. In wild-type HCT116 cells, *LIG3* is expressed cell-wide. In contrast, *LIG3*$^{-/-:mL3}$ cells showed a mitochondrial-exclusive expression pattern. In *LIG3*$^{-/-:mL3:nL3}$ cells, mL3 is detected in the cytoplasm and nL3 is over-expressed in the nucleus.

Figure 2.4

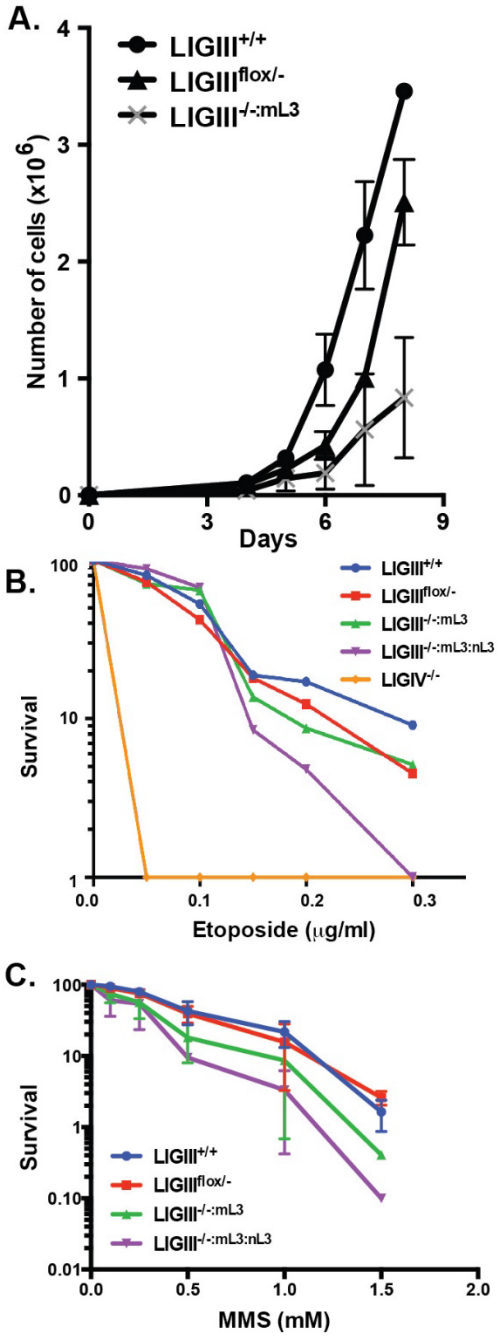

A.
LIGIII+/+
LIGIIIflox/-
LIGIII-/-:mL3
Number of cells (x10^6)
Days
B.
LIGIII+/+
LIGIIIflox/-
LIGIII-/-:mL3
LIGIII-/-:mL3:nL3
LIGIV-/-
Survival
Etoposide (μg/ml)
C.
LIGIII+/+
LIGIIIflox/-
LIGIII-/-:mL3
LIGIII-/-:mL3:nL3
Survival
MMS (mM)

Figure 2.4. A *LIG3* deficiency causes growth retardation, but not hypersensitivity to DNA damaging agents.

A) Three thousand cells were plated on day 0 and their growth was subsequently assessed by counting trypan blue-excluding cells at the indicated days. The average of two independent experiments, each done in triplicate, is shown. B) Etoposide sensitivity. Three hundred cells were plated a day before etoposide treatment. Survival was normalized by setting the value obtained from no etoposide treatment as 100 percent. As a positive control, a *LIG4*$^{-/-}$ cell line, which is known to be extremely sensitive to etoposide, was used. The plotted values are the average of three independent experiments. C) MMS sensitivity was performed in a similar way to the etoposide sensitivity test, except for the drug-treatment time: cells were incubated in MMS-containing media for 1 hr.

Figure 2.5

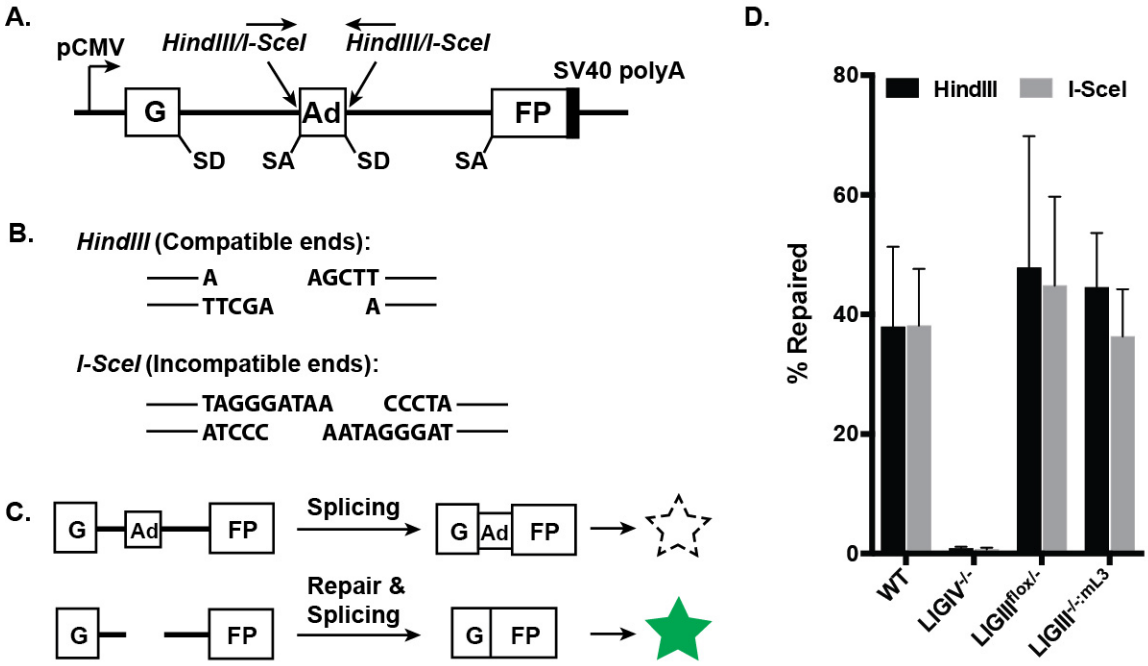


Figure 2.5. A *LIG3* deficiency does not affect total end-joining activity.

A) A cartoon of the pEGFP-Pem1-Ad2 repair substrate used for analysis of DNA end-joining activity. Expression of the GFP cassette is driven by the CMV promoter and terminated by the SV40 polyA sequence. The GFP coding sequence is interrupted by a 2.4 kb intron containing an adenovirus exon (Ad), which is flanked by *HindIII* and *I-SceI* restriction enzyme recognition sites. Splice donor (SD) and splice acceptor (SA) sites are also shown. B) *HindIII* digestion generates compatible ends with 4 nt overhangs, whereas *I-SceI* digestion produces incompatible ends, that require some processing before they can be rejoined. C) The starting substrate is GFP negative because the Ad exon is efficiently spliced into the middle of the GFP ORF, inactivating the GFP activity. Cleavage with either *HindIII* or *I-SceI* removes the Ad exon and, upon successful intracellular plasmid circularization, GFP expression is restored and can be quantitated by flow

cytometry. D) The impact of LIG3 deficiency on end joining. A *LIG4*-null cell line was used as a negative control. The plotted values are the average of three independent experiments.

Figure 2.6

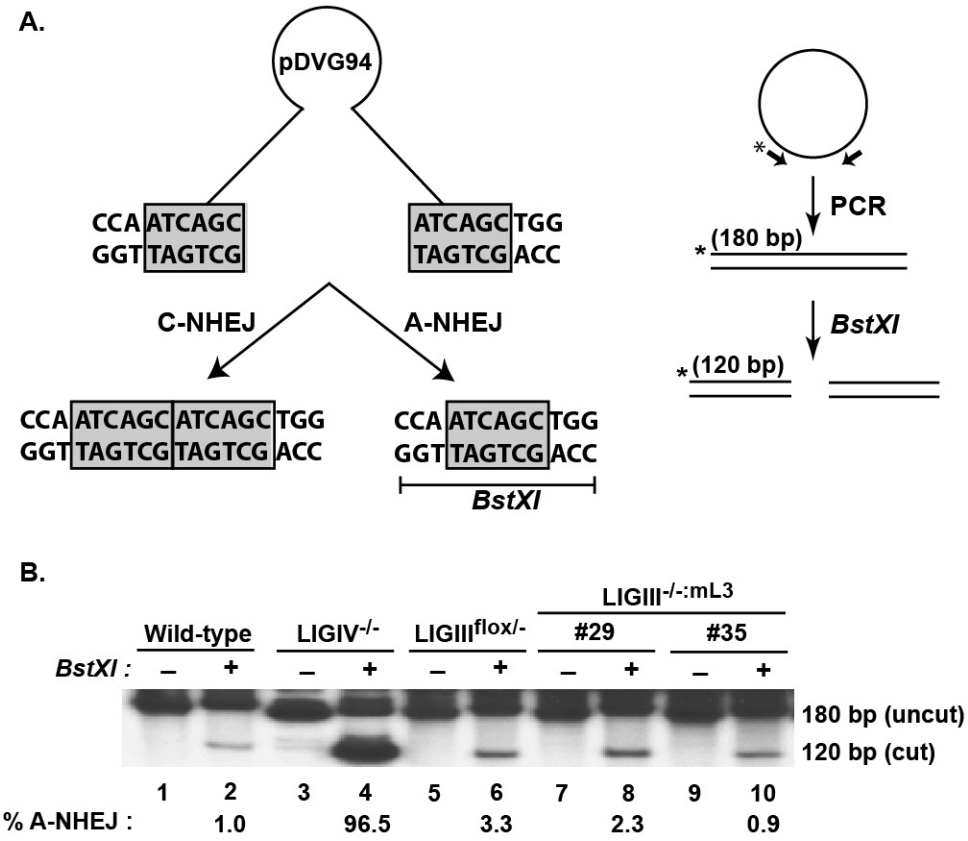


Figure 2.6. *LIG3*-null cells have normal microhomology-mediated end-joining activity.

A) After *EcoRV* and *AfeI* restriction enzyme digestion, the reporter substrate, pDVG94, becomes a blunt-ended linear plasmid with 6 bp direct repeats at both ends. Repair via C-NHEJ generally retains at least part of either repeat, whereas A-NHEJ generates only a single repeat and a novel restriction enzyme site that can be cleaved by *BstXI*. Repaired junctions were amplified by PCR using radiolabeled primers and the 180 bp PCR product was subjected to *BstXI* restriction enzyme digestion. The 180 bp uncut product represents repair via C-NHEJ whereas the 120 bp digested product represents A-NHEJ-mediated

repair. B) Both *LIG3* heterozygous (*LIG3*$^{flox/-}$) and LIG3-null (*LIG3*$^{-/-:mL3}$) cells have similar microhomology end-joining activity to the wild-type control. A *LIG4*$^{-/-}$ cell line was used as a positive control.

Figure 2.7

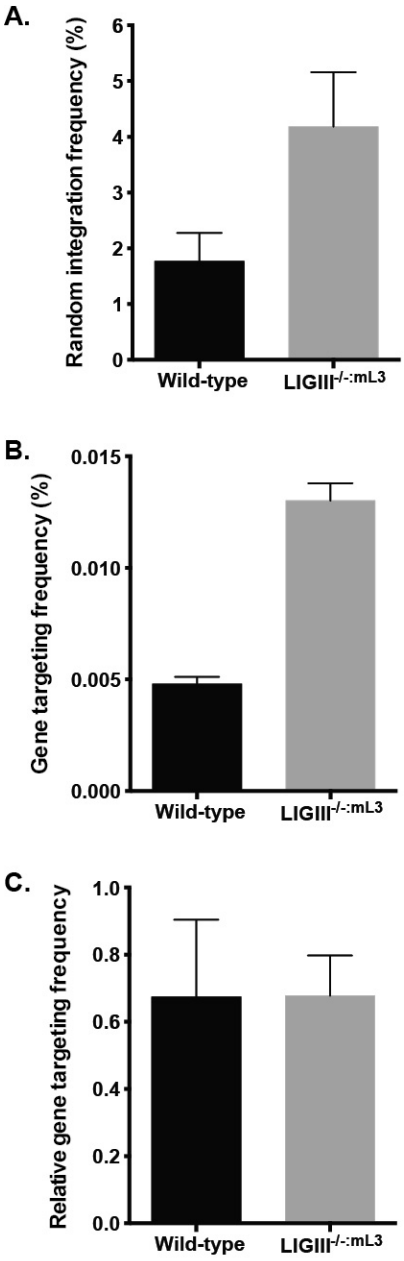

A.
Random integration frequency (%)
6
5
4
3
2
1
0
Wild-type
LIGIII-/-:mL3
B.
Gene targeting frequency (%)
0.015
0.010
0.005
0.000
Wild-type
LIGIII-/-:mL3
C.
Relative gene targeting frequency
1.0
0.8
0.6
0.4
0.2
0.0
Wild-type
LIGIII-/-:mL3

Figure 2.7. A LIG3 deficiency does not affect the relative frequency of rAAV-mediated gene targeting at the HPRT locus.

A) rAAV random integration frequency. The random integration frequency was determined as the number of puromycin-resistant clones normalized by plating efficiency. B) rAAV gene targeting frequency. The gene targeting frequency was determined as the number of both puromycin and 6-thioguanine-resistant clones normalized by plating efficiency. C) The relative frequency of rAAV gene targeting represents the ratio of the correct targeting events versus the total viral integration events. *LIG3*-null cells showed both higher random integration and gene targeting frequencies and hence overall there was no difference in relative gene targeting frequency.

Figure 2.8

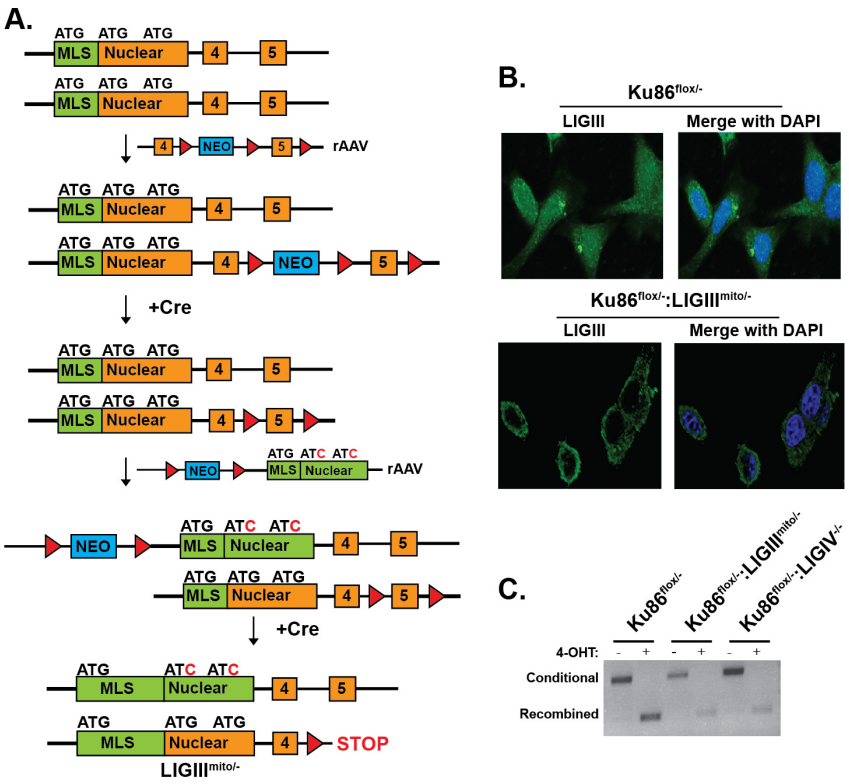
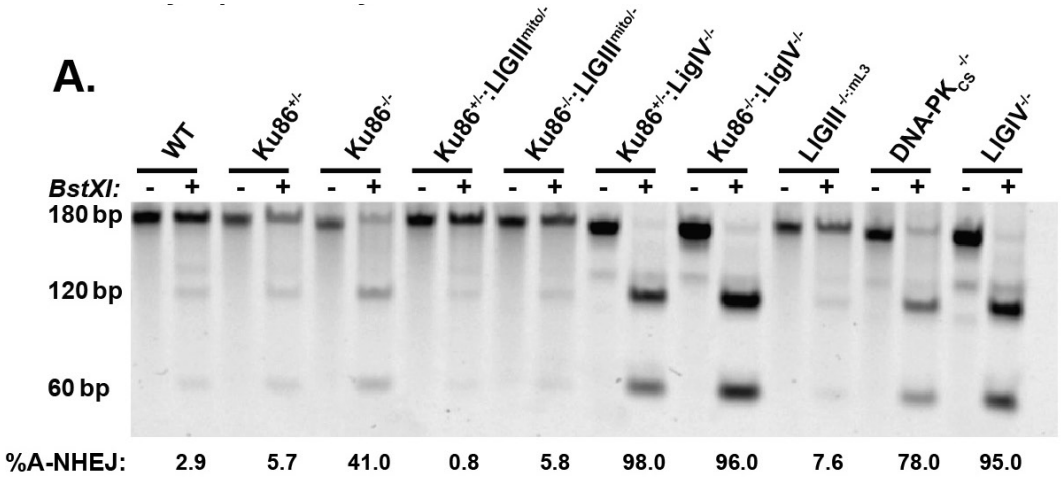


Figure 2.8 Scheme for the construction of a *KU86*:*LIG3* doubly-deficient human cell line.

A) A cartoon of part of the *LIG3* genomic locus showing the three relevant ATGs in exon 2 and the approximate location of the mitochondrial localization sequence (MLS, green rectangle) and the portion of LIII that encodes the nuclear isoform (Nuclear, orange rectangle). The numbered rectangles represent exons and lines in between are introns; exon 1 (non-coding), exon 3 and additional downstream exons are not shown. The conditional knockout vector had three loxP sites (filled triangles) that flanked a neomycin drug resistance gene (NEO, blue rectangle) and exon 5, respectively. After first round

targeting the NEO gene was removed by transiently treating the cells with Cre recombinase (+Cre). The second round of gene targeting utilized a vector into which the second and third ATGs had been mutated to ATCs. Following correct gene targeting and a repeated Cre treatment a LIG3$^{mito/-}$ cell line was constructed. B) Mitochondrial-exclusive localization of LIG3 in *LIG3*$^{mito/-}$ cells. In *KU86*$^{flox/-}$ cells, LIG3 was localized cell-wide. In contrast, *KU86*$^{flox/-}$:*LIG3*$^{mito/-}$ cells showed a mitochondrial-exclusive expression pattern. DAPI was used to identify the nuclei. C) PCR analysis was used to demonstrate that the addition (+) of 4-OHT induced efficient conversion of the *KU86*$^{flox/-}$ configuration to a recombined *KU86*$^{-/-}$ state in all 3 of the indicated cell lines.

Figure 2.9


A.
WT
Ku86+/-
Ku86-/-
Ku86+/-:LIGIIImito/-
Ku86-/-:LIGIIImito/-
Ku86+/-:LigIV-/-
Ku86-/-:LigIV-/-
LIGIII-/-:mL3
DNA-PKcs-/-
LIGIV-/-
BstXI: - + - + - + - + - + - + - + - + - + - +
180 bp
120 bp
60 bp
%A-NHEJ: 2.9 5.7 41.0 0.8 5.8 98.0 96.0 7.6 78.0 95.0

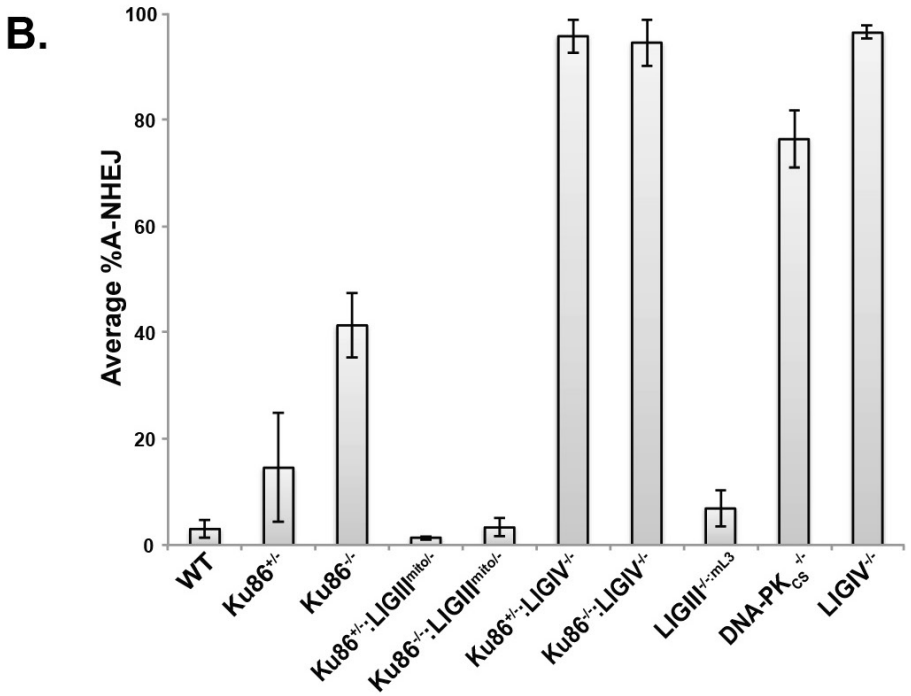

B.
Average %A-NHEJ
100
80
60
40
20
0
WT
Ku86+/-
Ku86-/-
Ku86+/-:LIGIIImito/-
Ku86-/-:LIGIIImito/-
Ku86+/-:LIGIV-/-
Ku86-/-:LIGIV-/-
LIGIII-/-:mL3
DNA-PKcs-/-
LIGIV-/-

Figure 2.9. LIG3 is required for microhomology-mediated A-NHEJ events normally repressed by Ku.

A) The reporter substrate, pDVG94, described in Figure 5 was used with the indicated cell lines. The 180 bp uncut product represents repair via C-NHEJ whereas the 120 bp digested product represents A-NHEJ-mediated repair. Note that the enhanced A-NHEJ repair observed in a *KU86*$^{-/-}$ cell line was completely abrogated in a *KU86*$^{-/-}$:*LIG3*$^{mito/-}$ cell line. B) Four independent assays comparable to that shown in panel (A) were averaged and the error bars represent the standard error of the mean.

Figure 2.10

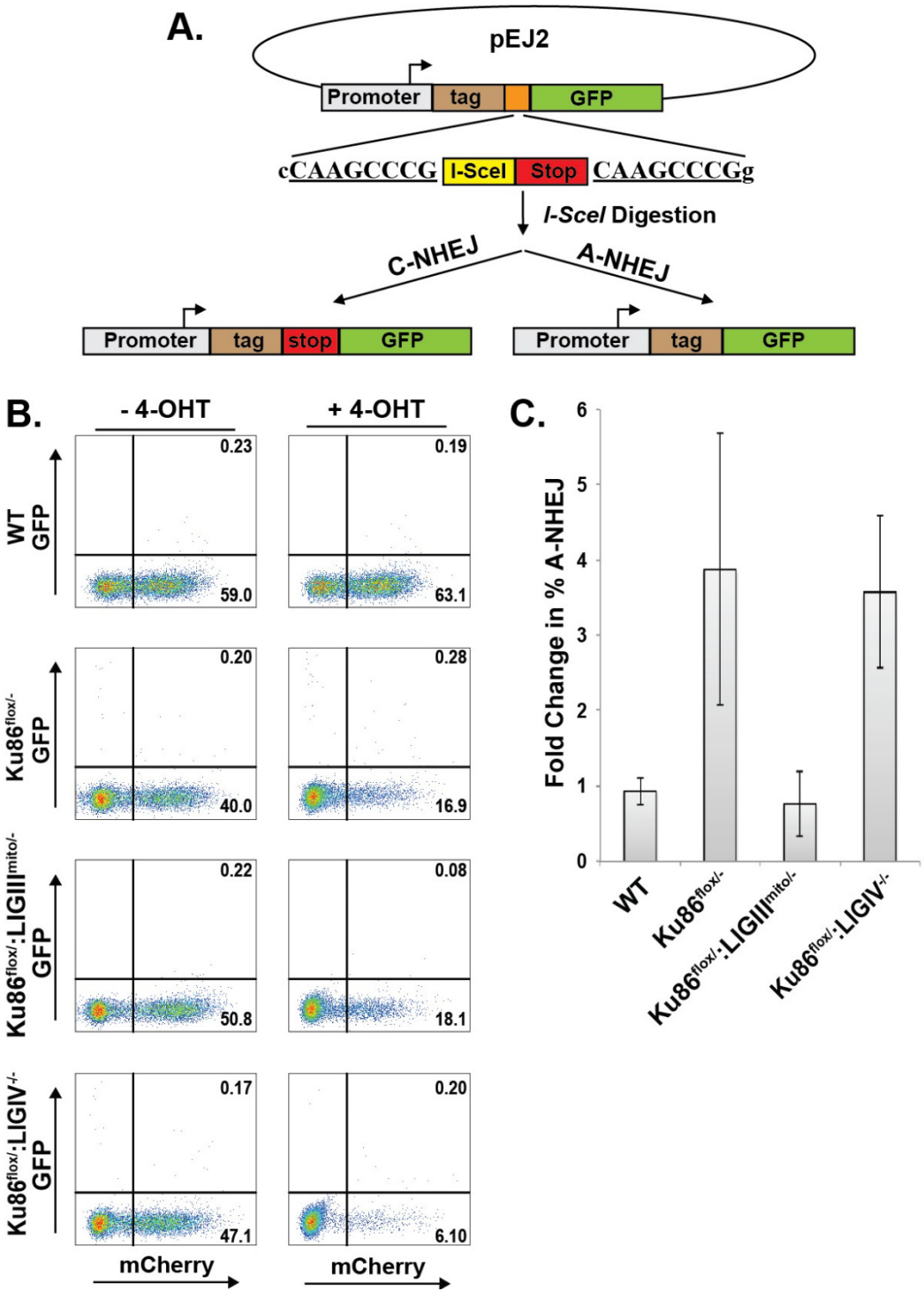

A.
pEJ2
Promoter
tag
GFP
cCAAGCCCG
I-SceI
Stop
CAAGCCCGg
I-SceI Digestion
C-NHEJ
A-NHEJ
Promoter
tag
stop
GFP
Promoter
tag
GFP
B.
- 4-OHT
+ 4-OHT
WT GFP
0.23
59.0
0.19
63.1
Ku86flox/- GFP
0.20
40.0
0.28
16.9
Ku86flox/-:LIGIIImito/- GFP
0.22
50.8
0.08
18.1
Ku86flox/-:LIGIV-/- GFP
0.17
47.1
0.20
6.10
mCherry
mCherry
C.
Fold Change in % A-NHEJ
0
1
2
3
4
5
6
WT
Ku86flox/-
Ku86flox/-:LIGIIImito/-
Ku86flox/-:LIGIV-/-

Figure 2.10. LIG3, and not LIG4, is required for microhomology-mediated EJing events normally repressed by Ku.

A) A diagram of the reporter substrate, pEJ2. The gray rectangle with the arrow represents a transcriptional promoter. The brown rectangle (tag) represents and in-frame epitope tag. The orange rectangle is expanded below to show an *I-SceI* restriction enzyme recognition site and the red rectangle represents translational stops in all 3 reading frames (stop). These elements are flanked by 8 bp of a direct repeat (CAAGCCCG) that permits microhomology-mediated repair. The green rectangle (GFP) represents the coding sequences for green fluorescent protein. Repair of the *I-SceI*-linearized pEJ2 by C-NHEJ retains the stop cassette and does not permit GFP expression whereas repair by A-NHEJ deletes out the stop element and allows GFP expression. B) The indicated cell lines were transfected with the *I-SceI*-linearized pEJ2 reporter and then either left untreated (-Cre) or treated with 4-OHT to induce Cre (+Cre) expression. C) Four independent experiments similar to the one shown in panel (B) were averaged and presented as the fold change relative to wild type.

Figure 2.11

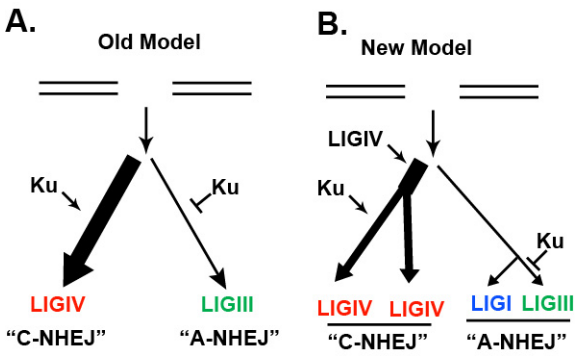


Figure 2.11. A new model for EJing pathways in human somatic cells

A) The old model for EJing. A DSB could be repaired by one of two pathways. The first and most active was the Ku-dependent, LIG4-dependent C-NHEJ pathway and the second was the minor Ku-suppressible, LIG3-dependent A-NHEJ pathway. B) A proposed new model posits that all DSB substrates are first subjected to a LIG4-mediated rejoining event, which is relatively successful. Only when this rejoining event fails is Ku recruited for facilitating additional processing events. LIG4 is, however, subsequently re-recruited to finish these repair events. Thus, C-NHEJ is actually composed of 2 separate pathways, one of which is Ku dependent and one of which is not. Both pathways are, however, dependent upon LIG4. Moreover, the A-NHEJ pathway also bifurcates and can either be mediated by LIG1 or LIG3 and only the later of which is Ku-suppressible.

Figure 2.12

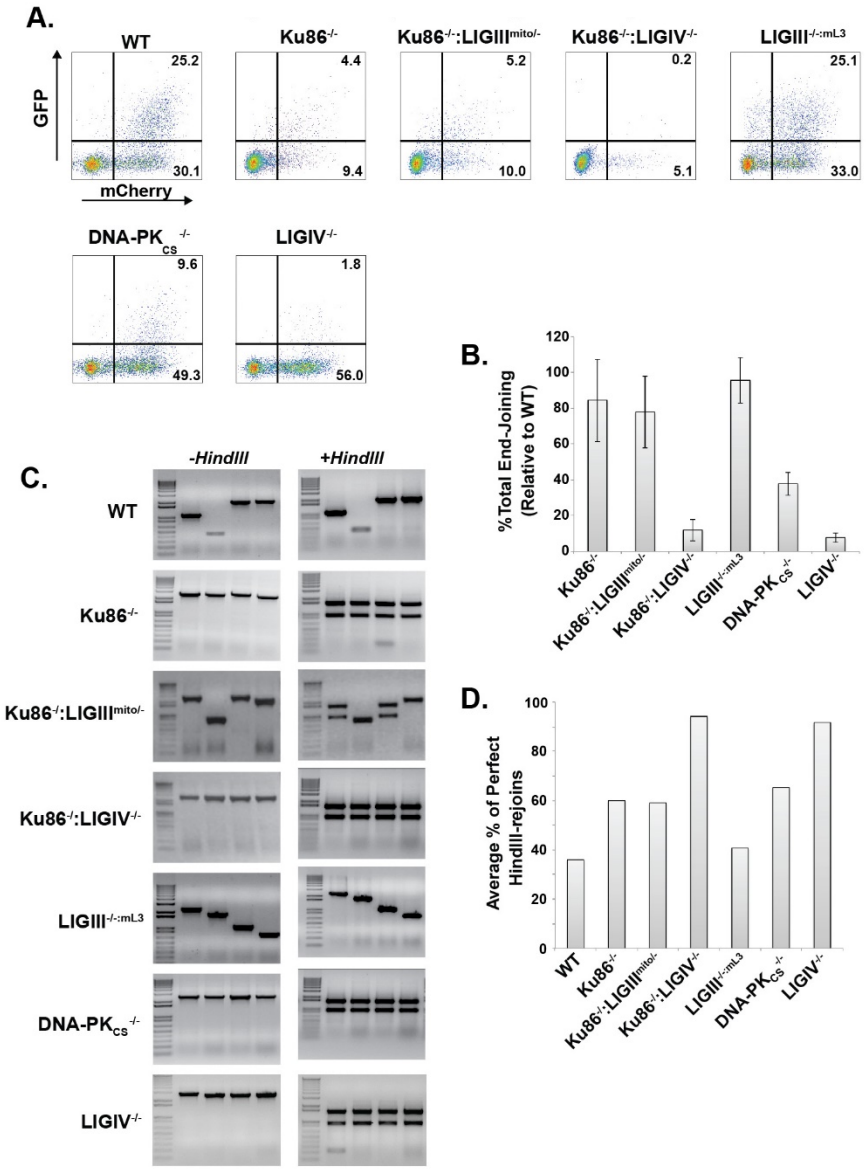


Figure 2.12. Ku-independent EJing is LIG4-dependent.

A) Profiles of EJing assays using the reporter pEGFP-Pem1-Ad2 described in Figure 4 transfected into the indicated cell lines. B) Four independent assays comparable to that

shown in panel (A) were averaged and the error bars represent the standard error of the mean. C) Repaired plasmids were recovered from the indicated cell lines, propagated thorough bacteria and then either left untreated (-*HindIII*) or subjected to restriction enzyme digestion (+*HindIII*) and then analyzed by agarose gel electrophoresis. The appearance of a diagnostic doublet following restriction enzyme digestion indicates that perfect rejoining occurred. All of the gels show a marker lane on the left followed by 4 lanes containing independently isolated plasmids. D) 18 to 40 independent plasmids similar to the four shown in panel (C) were quantitated for each of the indicated cell lines and scored for the occurrence of perfect joins.

Chapter 3

PARP1 is required for preserving genomic and telomeric integrity

Adam Harvey[1], Nicholas Mielke[1], Julia Grimstead[2], Thanh Nguyen[1], Mathew Mueller[1], Duncan M. Baird[2*] and Eric A. Hendrickson[1*]

[1]Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota Medical School, Minneapolis, MN 55455, USA

[2]Division of Cancer and Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff, CF14 4XN, United Kingdom

*Co-senior authors.
Correspondence can be addressed to: Eric A. Hendrickson, BMBB Department, University of Minnesota Medical School, 6-155 Jackson Hall, 321 Church St., SE. Minneapolis, MN 55455. E-mail: hendr064@umn.edu

Author Contributions:

- Harvey, A. was responsible for the experimental design of all the work shown. He performed the work presented in all figures with the following exceptions:
- Mielke, N. performed: Fig 3.1 C, D, E and Fig 3.3 A, B, (although he was assisted by Harvey, A.
- Mueller, M. performed Fig 3.8A and constructed the PARP1$^{-/-}$ cells.
- Ngyuen, T. performed work involved with Fig 3.8D; again, while being assisted by Harvey, A.
- Grimstead, J. performed work in Fig 3.10, under the direction of Baird, D.

Synopsis

PARP1 has been well studied and is clinically important because of its synthetic lethality with *BRCA1* and *BRCA2* mutations, which are causative for inherited breast and ovarian cancers. Biochemically, PARP1 is a single-stranded DNA break repair protein, is needed for preserving genomic integrity, and has been implicated in a bevy of additional cellular pathways and processes. Because of PARP1's influence on so many cellular pathways, its precise contribution to human cellular biology remains relatively obscure. To molecularly address this deficiency, we utilized gene editing to construct genetically-null *PARP1* human cancer cells. As expected, we found a specific role for *PARP1* in an alternative form of DSB repair. In addition, however, we also observed cell cycle progression defects and elevated endogenous DNA damage signaling. Moreover, we found evidence of telomere defects, which are likely the strongest contributor to the apparent genomic instability. Specifically, *PARP1*$^{-/-}$ cells have short telomeres that frequently co-localize with markers of endogenous DNA damage signaling. Our data suggest that while PARP1 does not participate significantly in DNA DSB repair itself, it does prevent the incidence of DSBs, particularly at telomeres, presumably by preventing replication-induced DNA damage.

Introduction

PARP1 is a ubiquitously and very abundantly expressed protein that post-translationally modifies target proteins with PAR moieties using NAD^+ as its source. Impressively, the sheer abundance of these post-translational modifications in cells enabled researchers to discover such modifications prior to any information about the protein(s) responsible for their catalysis (Hayaishi and Ueda 1977). Of the 17-known PARP-domain-containing proteins (named PARP1 through PARP17, respectively), PARP1 is the most ubiquitous and most active within eukaryotic cells, as its genetic deletion alone causes a dramatic loss in the amount of detectable PAR within cells (Bock and Chang 2016). Because PARP1 is the most abundant PARP and because PARylation is thought to be an important signaling process, it is not surprising that PARP1 has been implicated in a vast array of cellular processes, including cellular metabolism, cell cycle regulation, DNA replication, and DNA break repair (Mangerich and Burkle 2012). PARP1 has been studied biochemically *in vitro* and extensively by genetic knockout *in vivo* in model organisms including mice (Wang *et al.* 1995), plants (Boltz et al. 2014), and flies (Miwa *et al.* 1999), as well as in chicken DT-40 cells (Hochegger *et al.* 2006). These reports generally conclude that PARP1 is important to preserve genomic integrity, and that it primarily participates in the repair of SSBs (Helleday 2011). To date there has been no phenotypic characterization of the genetic knockout of *PARP1* in human somatic cells, as the field has either relied on RNAi knockdowns, or, primarily, by utilizing one of the many available inhibitors to PARP1 (Krishnakumar and Kraus 2010, Gibson and Kraus 2012). The use of RNAi, however, rarely completely eliminates PARP1 from a given cell, potentially obscuring relevant phenotypes. In a complementary fashion, PARP inhibitors

generally have a dominant-negative effect on cellular PARP1 by trapping PARP1 at a SSB in a DNA-bound state (Godon *et al.* 2008, Strom *et al.* 2011, Murai *et al.* 2012). Thus, the normal role of PARP1 in human cells remains somewhat poorly defined.

PARP1 is well-known because it exhibits synthetic lethality with *BRCA1* and *BRCA2* (Bryant et al. 2005, Farmer et al. 2005). A prevailing theory is that SSBs, which normally would be recognized for repair by PARP1, can accumulate over time such lesions would be converted to DSBs as a consequence of DNA replication (Helleday 2011). Because *BRCA1* and *BRCA2* are required for the HDR of these DSBs it has been postulated that it is this activity of *BRCA1*- and *BRCA2*-dependent repair that is required to preserve a viable level of genomic integrity (Rosen and Pishvaian 2014). This pathway, while clearly relevant to explain the impact of PARP1's absence on survival may, however, only be part of the story. For example, PARP1's association with the replication fork is required for Chk1-dependent activation in response to replication stress (Min *et al.* 2013). Therefore, the absence of PARP1 might also dysregulate the integrity of replication forks in addition to resulting in a higher frequency of lesions.

Besides impinging upon HDR, *PARP1* has been implicated in the repair of DSBs by actively regulating NHEJ. NHEJ involves the end-to-end ligation of two broken ends of double-stranded DNA, and can be sub-categorized as C-NHEJ or A-NHEJ pathways. The C-NHEJ pathway is dependent on the Ku70/86 heterodimer, a ring-shaped protein complex that binds the ends of broken DNA. The binding of the ubiquitously expressed and very abundant Ku heterodimer and subsequent activation of the C-NHEJ pathway can occur within seconds of a DSB occurring and is inherently repressive of A-NHEJ (Fattah et al. 2010, Shahar *et al.* 2012, Oh et al. 2014). Accordingly, A-NHEJ is thought to be a minor or back-up repair pathway in normal cells. The hallmark of A-NHEJ is the use of

microhomology, which also constitutes a molecular signature that remains at the site of repair, to facilitate the ligation of the two DNA ends (Ciccia and Elledge 2010). One documented pathological role for A-NHEJ is likely its involvement in oncogenic chromosomal translocations in mice (Simsek et al. 2011), although this is probably not the case in human somatic cells (Ghezraoui et al. 2014). *PARP1* has been implicated in the regulation of A-NHEJ, which was partially a consequence of discovering that PARP1-associated proteins, such as XRCC1, were required for A-NHEJ (Audebert et al. 2004, Wang et al. 2005, Wang et al. 2006, Audebert *et al.* 2008). Finally, it has been suggested that the DNA binding activity of PARP1 could compete with Ku to enable A-NHEJ to occur in place of C-NHEJ (Wang et al. 2006), or repress Ku's ability to access the DNA break (Hochegger et al. 2006).

Telomeres are the repetitive DNA:protein structures that serve to protect the ends of linear chromosomes from recognition as a DSB (Doksani and de Lange 2014). They are well known to regulate aging, as they gradually shorten over time due to the end replication problem. Moreover, and of important clinical significance, the activation of a telomere re-elongation pathway is a key requirement for malignant progression (Hanahan and Weinberg 2011). Telomeres appear to be difficult regions of the genome for DNA replication to occur (Verdun and Karlseder 2007), which is likely due to both the repetitive nature of telomeric DNA, combined with their tendency to form G-quadruplex structures (Smith and Feigon 1992, Parkinson *et al.* 2002). PARP1 has been identified as a telomere-binding protein and has been implicated in the regulation of telomere length maintenance (Gomez *et al.* 2006, Wang *et al.* 2007, Giannone *et al.* 2010, Salvati *et al.* 2010). Confusingly, the influence of *PARP1* loss-of-function in the mouse has been reported to result in telomere shortening (d'Adda di Fagagna *et al.* 1999) or to have no

impact what-so-ever (Samper *et al.* 2001). Similarly, there is a lack of agreement about the role of *PARP1* in telomere length maintenance in human cells as the use of either RNAi against *PARP1* (Beneke *et al.* 2008) or inhibitors to PARP1 (Beneke et al. 2008, Lee *et al.* 2015) resulted in either telomere shortening (Beneke et al. 2008) or lengthening (Lee et al. 2015). Since some of these findings seem mutually exclusive and are likely due to differences in the experimental systems employed, a role for *PARP1* in mammalian telomere maintenance is still controversial.

To experimentally address this issue, we utilized gene editing to generate *PARP1*-null human somatic cells. Human *PARP1*-null cells are viable and exhibit spontaneous DNA damage, which tends to localize at telomeres, and is co-incident with short telomeres. We further find that this is likely not due to a defect in DSB repair *per se*, but rather reflects an inability to mitigate damage associated with DNA replication.

Results

Creation of *PARP1*-null cells

We utilized gene targeting in human HCT116 cells to functionally inactivate *PARP1* by an exon-replacement strategy (Hirata *et al.* 2002). We designed a gene targeting construct in a rAAV, such that correct targeting would result in the replacement of *PARP1* exon 4 with a *NEO* drug selectable marker (Fig. 3.1A). Correct gene targeting was screened for by using PCR primer pairs with one primer that flanked the targeting construct combined with an internal primer specific to the drug selectable marker. After successful gene targeting, the drug selectable marker was removed by Cre-recombinase, thus generating a null allele. Two rounds of targeting were required to generate a diploid null cell line, which was confirmed using PCR primers that flank exon 4 (Fig. 3.1B; Fig. 3.2A). In a scenario where a gene exhibits no strong selective pressure, Mendelian genetics would predict that when targeting a heterozygotic cell line, there is an equivalent 50% chance of targeting either the already targeted allele ("re-targeting") or targeting the second, still functional allele. During the second round of *PARP1* targeting, only 3 of 72 correctly targeted clones resulted in the loss of the second *PARP1* allele (*i.e.*, 69 of 72 clones were re-targeted) (Table 3.1). This exceptional disequilibrium in the gene targeting frequency is usually a hallmark of genes that provide a significant growth disadvantage when absent (Dang *et al.* 2006, Fattah *et al.* 2008, Ruis *et al.* 2008). Thus, although isolation of three independent *PARP1*-null clones was unequivocal evidence that *PARP1* is not essential in human HCT116 cells, the frequency with which these clones were obtained was also a first indication that *PARP1* had an important role in human cellular biology.

We next sought to complement these cells with either an empty vector, or a wild type (WT) *PARP1* cDNA, which was integrated randomly into the genome by a PiggyBac transposon system (Doherty *et al.* 2012). The restoration of PARP1 protein in the null cells was confirmed by western blot (Fig. 3.1C; Fig. 3.2C). A series of complemented clones in which PARP1 was either under- ("low"), ~normally- ("medium") or over- ("high") expressed were generated for each of the three null cell lines (Fig. 3.1c; Fig. 3.2C). In order to validate that the complemented clones contained active protein, a PARP-activity assay was performed, which confirmed both the successful ablation of PARP activity in the null cells, as well as their functional complementation (Fig. 3.1D). Interestingly, even the lowest levels of complementing PARP1 expression could fully rescue PARP activity (Fig. 3.1d). An obvious phenotype of the *PARP1*$^{-/-}$ cells was their slow growth, as they exhibited an almost 50% reduction in doubling time (Fig. 3.1E; Fig. 3.2B). Again, this phenotype could be completely rescued by the re-expression of even low levels of PARP1 (Fig. 3.1E). In summary, these data compellingly demonstrated that *PARP1* is not essential in human somatic cells, but that its absence results in significant deficits to both proliferation and survival.

PARP1-null cells accumulate in the G_2 phase of the cell cycle

To better understand the cellular growth defect of *PARP1*-null cells, we investigated whether it was correlated with deficits in cell cycle progression. In asynchronous populations, the null cells exhibited a modest increase in the number of cells in G_2, compared to both WT, and the complemented cells when their DNA content was analyzed (Fig. 3.3A). In order to better understand this G_2 accumulation, cells were synchronized at G_1/S with serum starvation, followed by an overnight incubation in thymidine (in the

presence of serum), which transiently arrested the cells at the G_1/S transition. After releasing the cells into standard medium, the cell cycle profile of the cells was determined (Fig. 3.3B). The *PARP1*-null cells progressed through the cell cycle at approximately the same rate as either WT or complemented cells, but after S-phase, dramatically accumulated in the G_2 phase. This is best exemplified by the amount of *PARP1*-null cells remaining in G_2 (36.4%) at the 12 hr time point, in comparison to the rest of the genotypes (12% to 18%), which were successfully able to continue through mitosis and into the subsequent G_1 phase (Fig. 3.3B). The accumulation of *PARP1*-null cells in G_2, suggested that the cells were experiencing an elevated level of DNA damage that was, in turn, activating the G_2/M checkpoint. Consistent with this interpretation, *PARP1*-null cells had elevated levels of p53, with respect to control lines (Fig. 3.3C; Fig. 3.4).

PARP1 modulates, but is not required for A-NHEJ

The spontaneous elevation of p53 expression in asynchronously growing *PARP1*-null cells suggested that the absence of *PARP1* was contributing to an accumulation of DNA damage. Given the aforementioned reports of *PARP1*'s role in DSB repair, we next measured the cells' capacity for this activity. Cells were transfected with a linearized plasmid-reporter, pDVG94, allowed 48 hr to enact repair, and then circularized plasmids were recovered from the transfected cells. Cells have two options to repair the linearized plasmid. They can simply re-ligate the ends together, which is indicative of C-NHEJ and which can be quantitated as a ~180 bp PCR product when primers flanking the repair junction are utilized (Fig. 3.5A). Alternatively, cells which utilize the 6 bp of microhomology present at the linearized ends to repair the plasmid create a diagnostic restriction enzyme recognition site for *BstxI*. Cleavage of the PCR products generated with primers flanking

the repair junction with *BstXI* generates a 120 bp fragment (and a 60 bp fragment) and the appearance of this product versus the 180 bp product enables a relative measure of A-NHEJ versus C-NHEJ activity (Fig. 3.5A) (Verkaik et al. 2002). In wild type cells approximately 15% of the repair products could be ascribed to A-NHEJ (Fig. 3.5B, C). As a positive control we also analyzed the same plasmid rejoining in a cell line defective in DNA Ligase 4 (*LIG4*) (Oh et al. 2013). As expected (Fattah et al. 2010, Oh et al. 2014) these cells carried out virtually exclusively (97%) A-NHEJ (Fig. 3.5B, C; Fig. 3.6). Consistent with previous studies (Audebert et al. 2004, Wang et al. 2005, Wang et al. 2006, Audebert et al. 2008), the *PARP1*-null cells showed a statistically significant deficit in A-NHEJ (Fig. 3.5B, C). The deficit, however, was rather small (<2-fold), and importantly could not be phenocopied by treating wild type cells with a PARP1 inhibitor, olaparib (Fig. 3.6). To clarify these results, we next induced A-NHEJ activity by pretreating cells with the DNA-PK_{cs} inhibitor, NU7441 (Leahy *et al.* 2004), which should increase the relative amount of A-NHEJ by inhibiting the Ku/DNA-PKcs-dependent C-NHEJ pathway. All genotypes, except as anticipated, LIG4-null cells, showed enhanced A-NHEJ activity in the presence of NU7441 (Fig. 3.5B, C). Importantly, however, the *PARP1*-null cells showed increases in A-NHEJ activity comparable to the wild-type and complemented clones (Fig. 3.5B, C). Most provocatively, the treatment of *LIG4*- and *DNA-PK_{cs}*-null clones with olaparib was completely ineffective in inhibiting A-NHEJ (Fig. 3.6). Thus, the absence of *PARP1* did not affect the cellular capacity for A-NHEJ in all situations where A-NHEJ activity was either genetically or chemically enhanced.

Since previous models had suggested that *PARP1* may compete for DSBs with the Ku heterodimer (Hochegger et al. 2006, Wang et al. 2006) we carried out an additional experiment to test whether or not the converse of our conclusion that *PARP1* activity might

be C-NHEJ dependent was true: *i.e*., whether the absence of *PARP1* affects C-NHEJ. *PARP1*$^{-/-}$ cells and relevant controls were transfected with the pEGFP-Pem1-Ad2 reporter of C-NHEJ activity (Seluanov *et al.* 2004). As expected, a *LIG4*-null cell line showed greatly reduced activity in this assay (Fig. 3.5D; Fig. 3.7). In contrast, the presence or absence of *PARP1* had no effect on the levels of C-NHEJ in the various cell lines (Fig. 3.5D; Fig. 3.7). Thus, we conclude that *PARP1* does not participate (significantly) in either C-NHEJ- or A-NHEJ-mediated DSB repair in human cells.

PARP1 is required for proper telomere maintenance

The above experiments demonstrated that although *PARP1*-null cells seemed to be sensing significant amounts of DNA damage or stress, they had only mild deficits in the repair of such damage. Thus, we hypothesized that this damage might be associated with stalled or stressed DNA replication forks. One region of particular interest was telomeres (Janson *et al.* 2015), as PARP1 had been identified as a telomere-binding protein, and some previous reports of telomere shortening have been associated with PARP1 inhibition or inactivation (Gomez et al. 2006, Wang et al. 2007, Giannone et al. 2010, Salvati et al. 2010). To explore this possibility, we analyzed metaphase spreads from wild type and *PARP1*$^{-/-}$ cells. *PARP1*$^{-/-}$ metaphases were frequently distorted, a hallmark of genomic instability, and exhibited a significant increase in the amount of signal-free ends, after labeling the telomeres with a PNA-probe (Fig. 3.8A, B). Thus, *PARP1*$^{-/-}$ cells appeared to have at least a subset of very short telomeres. Deficits in telomere length were confirmed by Southern blotting. Initial screening of several subclones of WT, *PARP1*$^{-/-}$, and *PARP1*$^{-/-:+PARP1}$ complemented cells demonstrated that the null cells had significantly shorter telomeres than WT cells (Fig. 3.8C). We did note that this phenotype was variable, *i.e.*

some null clones were shorter than others, and not all of the complemented clones restored the telomere length to WT levels. We attribute this to the fact that telomere length can be variable on a clonal basis in telomerase-immortalized cancer cells. Thus, to properly analyze the capability of PARP1 expression to complement the *PARP1*$^{-/-}$ cells, we created more independent complemented clones, along with empty vector controls, and re-screened their telomere lengths. Whereas none of the empty vector (EV)-containing clones complemented telomere length, the majority of the PARP1-expressing clones did restore telomere length, albeit usually not to wild type lengths (Fig. 3.8D). Thus, we conclude that PARP1 prevents abnormal telomere shortening, but it does not significantly contribute to telomere lengthening.

Given the aforementioned observed G_2/M cell cycle arrest, we probed for any connection between the spontaneous DNA damage and the shortened telomere phenotype. We utilized an IF-FISH hybrid assay, which combines immunoflouresence of proteins with FISH to co-visualize proteins and DNA sequences. We implemented this assay in the various cell lines for both telomeric DNA and 53BP1, a common marker of DNA DSBs (Huyen *et al.* 2004, Dimitrova *et al.* 2008). In *PARP1*-null cells there was a higher spontaneous frequency of 53BP1 foci (Fig. 3.9A, C), and these foci tended to co-localize (telomere dysfunction-induced foci; TIFs) with telomeric DNA (Fig. 3.9A, B). Importantly both the elevated 53BP1 foci and the TIFs could be completely suppressed by the re-expression of PARP1 in the *PARP1*-null cells (Fig. 3.9). Together, these data demonstrated that the telomeres of *PARP1*-null cells are short and that they are prone to incurring significant amounts of DNA damage.

PARP1 affects cellular immortalization

To confirm and extend the conclusion that *PARP1*-null cells have dysfunctional telomeres, we next tested whether *PARP1*-null cells could survive a "telomere challenge". In these experiments, a dominant negative telomerase (DN-h*TERT*) is expressed in cells to shorten their telomeres and force them into a "crisis" that is very much akin to the crisis that primary cells encounter as they evade senescence on their way to cellular immortality (Hendrickson and Baird 2015). Following the replicative erosion induced by the expression of DN-hTERT, cells normally undergo a period of slow growth and genetic instability due to the resulting shorten telomeres, but ultimately re-establish wild-type telomerase expression and telomere maintenance (Jones *et al.* 2014, Liddiard *et al.* 2016). Indeed, when DN-h*TERT* was expressed in wild type cells, all clones (15/15) analyzed escaped the subsequent crisis and continued to proliferate for at least 80 days, (or in some cases 120 days), after which point the experiment was intentionally terminated (Fig. 3.10). In stark contrast, only 1 of 10 *PARP1*-null clones was able to escape and immortalize (Fig. 3.10). These data demonstrated that the telomere dysfunction observed in *PARP1*-null cells severely compromised their ability to re-establish telomere maintenance in the face of gradual telomere erosion.

Discussion

PARP1 has been the subject of intense study in multiple model organisms. In spite of this, the molecular mechanism of PARP1 action in certain cellular transactions is still unclear. For example, whereas *PARP1* loss-of-function mutations were initially discovered to be synthetically lethal with *BRCA1* and *BRCA2* mutations (Bryant et al. 2005), some cancer patients with such mutations have not benefited from PARP1 inhibition (Fong et al. 2009), and many tumors that are *BRCA1*- and *BRCA2*-proficient can likewise be sensitized by PARP1 inhibition (Gelmon *et al.* 2010, Garnett *et al.* 2012). Thus, although there is no dispute that PARP1 inhibition can cause synthetic lethality in *BRCA1* and *BRCA2* mutant tumors, the molecular mechanism of that lethality is yet to be fully elucidated. One explanation for this ambiguity is that a common feature of PARP1 inhibitors is that they can have dominant-negative effects in cell lines that contain the target protein (Godon et al. 2008). While the results drawn from those experiments are not in any way invalid, it can be difficult to discern the effect of non-functional protein versus the absence of that protein. Here, we genetically ablated *PARP1* in a *BRCA1*- and *BRCA2*-positive cancer cell line, HCT116, to better understand the role of *PARP1* in an otherwise normal background. Unexpectedly, we find little evidence of significant DSB repair defects, but observe an increase in cells accumulating at the G_2/M checkpoint, likely resulting from an accumulation of DSBs. We further show that these DSBs tend to occur in telomeric DNA, which contribute to checkpoint activation, and further limit the cell's proliferation and subsequent ability to handle telomere stress.

One novel finding from these studies is the demonstration that *PARP1* is a non-essential gene in transformed human somatic cells. Our ability to isolate three

independent *PARP1*-null clones is unequivocal evidence that PARP1 is not required for survival. With that said, there has never been, to our knowledge, a human patient described anywhere in the world who is/was *PARP1*-null. Indeed, our own gene targeting studies argue strongly that PARP1, although not technically essential, is nonetheless so important that the development of a viable human is unlikely. Gene targeting is a completely egalitarian process and either allele in a diploid cell is as equally likely to be modified as the other (Hendrickson 2008). During the second round of *PARP1* targeting however, 69 of 72 clones were re-targeted and only 3 of 72 correctly targeted clones resulted in the loss of the second *PARP1* allele (Table 5.1). This exceptional disequilibrium in the gene targeting frequency is a hallmark of genes that provide a significant growth disadvantage when absent (Dang et al. 2006, Fattah et al. 2008, Ruis et al. 2008). This finding was unexpected, as PARP1-null mice are viable and fertile(Wang et al. 1995, de Murcia et al. 1997), and suggests a difference between PARP1 in humans and mice. Our subsequent demonstration that there are significant deficits in *PARP1*-null cells with telomere maintenance are completely consistent with this conclusion. Thus, the three independent *PARP1*-null clones notwithstanding, we predict that in the context of the embryonic development in humans that *PARP1* will be essential.

An additional key finding we present is the lack of a significant effect on A-NHEJ caused by the absence of *PARP1* (Fig. 3.5; Fig. 3.6). While it is compelling that PARP1 inhibition can result in diminished end-joining activity in many cell types, we suggest that *PARP1* is not a critical A-NHEJ gene. Such mischaracterization has historical precedent, as *PARP1* has been previously mislabeled as a core BER gene (Helleday 2011). This was originally suggested by the finding that *PARP1*$^{-/-}$ MEFs were hypersensitive to BER-sensitizing alkylating agents, such as methyl-methane sulfonate (Dantzer et al. 1999).

However, subsequent investigation demonstrated that the inhibition of PARP1 was simply either trapping a BER-intermediate (Strom et al. 2011) or modulating the terminal ligation step (de Murcia and Menissier de Murcia 1994) and that PARP1 was not an integral component of the BER machinery. We hypothesize that such a similar scenario exists for A-NHEJ and PARP1. The lack of a strong A-NHEJ phenotype for *PARP1*$^{-/-}$ human cells is consistent with emerging molecular evidence detailing the mechanism of A-NHEJ itself. Rather than a discrete subpathway of NHEJ, it now appears that A-NHEJ may rather be a HDR subpathway which is engaged when canonical HDR fails to find the appropriate homologous template for proper repair. Several independent reports support this emerging theory. First, conceptually, the requirement for microhomology at the ligation junction underlies a requirement for some degree of DNA-resection, followed by homology searching, both of which are much more akin to HDR than NHEJ. Moreover, reports have confirmed that the homology searching in A-NHEJ is dependent on the MRN complex and CtIP (Bennardo et al. 2008, Rass et al. 2009), as well as BRCA1 (Badie et al. 2015), all of which are canonical HDR genes. In addition, the kinetics of A-NHEJ are similar to HDR and distinct from C-NHEJ (Wang et al. 2001). *In toto*, these reports are consistent with A-NHEJ being a subpathway of HDR. If this model is true, then the key regulatory A-NHEJ genes are more than likely to be the upstream HDR repair genes, rather than *PARP1*. Thus, we suggest that while PARP1 inhibition could affect A-NHEJ activity in certain experimental models, it is not a canonical A-NHEJ gene. With this said, a small, albeit significant and reproducible, deficit in A-NHEJ activity was observed in PARP1-null cells — intriguingly however, only when they were proficient for DNA-PK$_{cs}$ (and therefore presumably C-NHEJ) (Fig. 3.5, Fig. 3.6). This deficit is more compatible with the more widely accepted models of A-NHEJ being a salvage pathway for ineffective C-NHEJ

(Iliakis 2009). Needless to say, these models are not mutually exclusive and A-NHEJ could be the salvage pathway for both unsuccessful HDR and C-NHEJ. In this scenario, the presence (or absence) of PARP1 seems to impact the C-NHEJ salvage subpathway more than the HDR one. All of these models clearly deserve further experimentation.

The role of PARP1 in telomere maintenance has remained an ambiguous, yet intriguing, concept. To date, the majority of work has described a role for PARP1 in mediating aberrant DNA repair at uncapped or damage telomeres, specifically causing the fusion of sister-chromatid telomeres (Salvati et al. 2010, Sfeir and de Lange 2012, Badie et al. 2015, Rai *et al.* 2016). Thus, these reports have implied that PARP1 is actively repressed from binding to functional telomeric DNA. Yet, other reports have indicated a functional interaction withTRF2, a principal component of the Shelterin complex (Gomez et al. 2006). PARP1 was also independently identified as a Shelterin binding protein by an unbiased mass-spectroscopy approach (Giannone et al. 2010). Consistent with those reports is the fact that PARP1 possesses a canonical TRF2-interacting motif (F/Y-X-L-X-P): $_{737}$YTLIP$_{741}$ (Chen et al. 2008). Reports of the role of *PARP1* in MEFs are conflicting: certain *PARP1*$^{-/-}$ MEFs exhibit telomere shortening (Gomez et al. 2006, Wang et al. 2007), whereas in other studies reported there was no appreciable telomere phenotype (Samper et al. 2001). Our data strongly suggest that one of the critical roles of *PARP1* in human somatic cells is to maintain telomeric integrity. The most likely scenario is that PARP1 is preferentially recruited to telomeres, through its interaction with Shelterin, to help regulate the repair of SSBs caused, or encountered by, the DNA replication machinery. Thus, the absence of PARP1 could result in the conversion of these telomeric SSBs to DSBs by DNA replication, resulting in a telomere shortening phenotype, DNA DSB signaling, and genomic instability — all of which we observed (Figs. 3.8 and 3.9). Importantly, we do not

suggest that PARP1 is a telomere lengthening protein; it does not function akin to telomerase and the re-introduction of *PARP1* to *PARP1*-null cells did not result in extensive telomere elongation. Rather, we posit that the presence of PARP1 allows for longer telomeres to maintain their stability. This is evidenced by the variation we observed in the extent of the telomere length restoration in *PARP1*-null complemented cells. The absence of PARP1 does cause telomere shortening (albeit indirectly), but the re-expression of PARP1 in these cells only allowed cells to stabilize the longer telomeres that were subsequently generated by the clonal variation in telomerase-positive cancer cells (Fig. 3.8). Thus, we conclude that PARP1's primary role is in preserving telomere length.

Recently, our laboratories have examined the contributions of the C-NHEJ and A-NHEJ pathways in facilitating the fusion of short dysfunctional telomeres in human cells following replicative erosion (Jones et al. 2014, Liddiard et al. 2016). These studies identified *LIG 3* as being essential for cells to escape the subsequent crisis and survive (Jones et al. 2014). Here we utilized this assay to confirm that *PARP1*-null cells have a role in telomere length maintenance. Thus, the expression of DN-hTERT in cells shortens their telomeres and forces them into a telomere crisis. In order to survive this crisis, cells must somehow re-establish telomere lengthening and then stabilize these new telomeres. The near inability of *PARP1*-null cells to do this (Fig. 3.10) is completely consistent with our posited role for PARP1 in telomere length maintenance.

Finally, we note that our data have clinical implications. Thus, PARP1 inhibitors are currently being extensively utilized in the clinic. Our demonstration here that the absence of PARP1 in human cells leads to aberrant telomere maintenance suggests that there may

be significant long-term repercussions to the chemical inhibition of PARP1 in human cells that might not be immediately evident.

Materials and Methods

Cell culture

HCT116 cells were purchased from the ATCC and maintained in McCoy's 5A media supplemented with 10% FBS, 1% glutamine, and 1% penicillin/streptomycin. Cells were maintained in 10 cm plates and passaged every 3 to 5 days. To initiate telomere erosion in HCT116 cells expressing DN-h*TERT*, the cells were transduced with amphotropic retroviral vectors containing a DN-h*TERT* cDNA (Hahn *et al.* 1999) as described (Preto *et al.* 2004). For cell synchronization studies, cells were cultured for 16 hr in McCoy's 5A media containing 0.1% FBS, and subsequently grown in 2 mM thymidine for 24 hr. Cells were then released into complete McCoy's 5A media and collected by trypsinization at the indicated times.

Gene targeting and *PARP1* knockouts

The PARP1 gene knockout by exon replacement with rAAV was performed by rAAV-mediated gene targeting. Briefly, homology arms were constructed by PCR, flanked by a LoxP-IRES-Neo-LoxP cassette, and ligated into an rAAV production vector. Producer 293-AAV cells were co-transfected with pAAV Helper and pAAV Rep/Cap, as described (Khan *et al.* 2011). Target wild type HCT116 cells (1×10^5) were plated approximately 24 hr prior to rAAV-infection in a 6-well plate. Cells were infected with virus-containing media, and 48 hr-post infection, the cells were single-cell subcloned in the presence of 0.5 mg/mL G418. Drug-resistant colonies were collected ~2 weeks after infection, and the correct replacement of exon 4 was screened by PCR. Correctly targeted clones were plated (1×10^5) and infected with an adenoviral vector expressing the Cre-recombinase to remove

the drug selectable marker by Cre-recombination. Cells were again single cell sub-cloned, and screened for correct Cre recombination events by PCR flanking exon 4. This process was repeated stepwise to remove the second PARP1 allele.

DNA repair assays

All transfections were performed on 5 x 10^5 cells with Lipofectamine 3000 in 6-well plates, which had been subcultured 24 hr prior to transfection. For the A-NHEJ reporter assay, we transfected 2.5 μg of linearized pDVG94 into target cells and allowed 24 hr for repair. The cells were subsequently collected by trypsinization, and re-circularized plasmids were recovered using conventional small-scale plasmid DNA isolation, as proper repair of the linearized junction by human cells creates a circularized DNA product which is accordingly recoverable. Repaired DNA junctions were PCR amplified using the FM30 and DAR5 primers (Verkaik et al. 2002). PCR products were then digested with the *BstXI* restriction enzyme. Digested PCR products were resolved by electrophoresis on a 6% polyacrylamide gel. The gel was then stained with SybrGold and imaged on a Typhoon FLA 9500 imager.

For the FACS-based NHEJ reporter assay, we first subcloned the ISce-I coding sequences from an expression plasmid (Jasin 1996) and added a C-terminal T2A-mCherry epitope by fusion PCR. We then cloned this expression construct into a pcDNA 3.1 expression vector. For each NHEJ FACS assay, 1.25 μg of pGEM-Ad2-EGFP was co-transfected with the ISceI-T2A-mChery plasmid into 5 x 10^5 cells in 6-well plates. 24 hr following transfection, cells were collected by trypsinization, fixed with 4% formaldehyde, and subjected to FACS analysis.

Telomere/terminal restriction fragment (TRF) assay

Genomic DNA was extracted from ~1 x 10^7 cells, and 50 μg of genomic DNA was digested with *HinfI* and *RsaI*, as described (Henson *et al.* 2009). For each sample, 12 μg of digested genomic DNA was resolved overnight on a 0.7% agarose 1 x TBE gel. This gel was depurinated, denatured, and neutralized, followed by overnight capillary transfer to a nitrocellulose membrane. The membranes were pre-hybridized for 1 hr with Church's buffer, then hybridized with a γ-P^{32}-endlabeled telomere probe in 4X SSC at 55°C overnight. Membranes were washed 3 times with 4X SSC and once with 4X SSC + 0.1% SDS, each for 30 min, exposed to a phosphorimaging screen, and detected and quantitated with a Typhoon phosphoimager.

Immunofluorescence and telomere FISH (IF-FISH)

This assay was performed as described (Conomos et al. 2014, Napier *et al.* 2015). Briefly, cells (1 x 10^5) were plated on chamber slides, and allowed to grow for 24 hr. Cells were washed once with PBS, then fixed with 4% formaldehyde in 1X PBS. Blocking and RNaseA treatment (0.1 mg/mL) were performed in antibody dilution media (ABDIL; 20 mM Tris-HCl pH 7.4, 0.2% fish gelatin, 2% BSA, 0.1% Triton X-100, 150 mM NaCl, and 0.1% sodium azide) at room temperature for 30 min. Cells were stained with a 53BP1 antibody (Ab36823), which was diluted in ABDIL for 1 hr, washed 3 times with 1X PBS + 0.1%Tween-20 (PBST), and incubated with an Alexa-488 goat IgG secondary antibody diluted in ABDIL for 1 hr. Cells were washed in PBST, fixed with 4% formaldehyde and prepared for FISH hybridization. A Telo-C PNA probe was hybridized to the slides at 80° in hybridization buffer (10 mM Tris-HCl pH 7.4, 4 mM Na_2HPO_4, 0.5 mM citric acid, 1.25

mM $MgCl_2$, 0.25% blocking reagent and 70% formamide). Slides were washed, counterstained with DAPI, and mounted with ProLong Gold (ThermoFisher). Microscopy was performed with a Nikon-TiE deconvolution bright-field microscope with a 60X objective.

Tables

Table 3.1. PARP1 gene targeting results

Desired Genotype	Targeted/Random Insertion	Number of Targeted Clones	Expected Number of Desired Clones
$Parp1^{+/-}$	23/96	23	23
$Parp1^{-/-}$	72/139	3	36

Figures

Figure 3.1

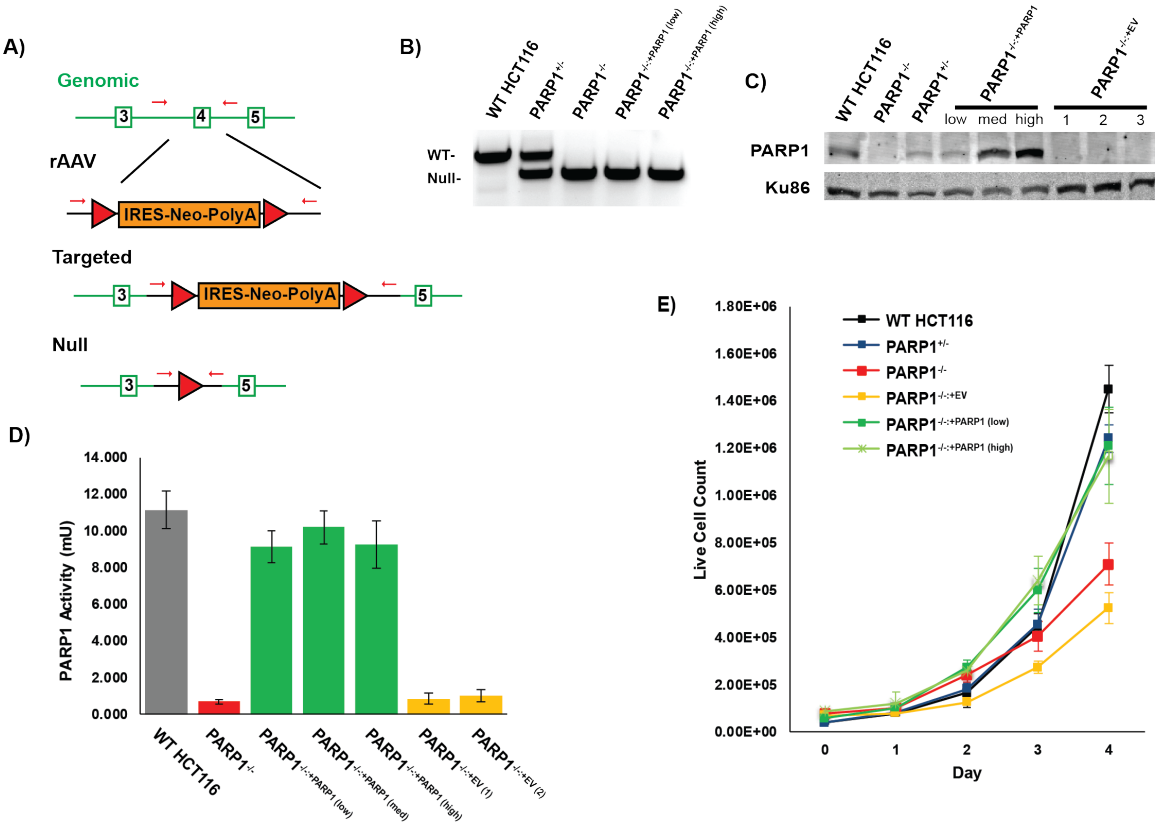


Figure 3.1. Construction and confirmation of *PARP1*-null cells

A) *PARP1* knockout HCT116 cells were constructed by rAAV-mediated gene targeting. Exon replacement of the 4th exon (open green rectangle) of the *PARP1* gene with a floxed, Neo-cassette (orange rectangle) occurs by HDR, which can be subsequently removed by Cre-recombinase to result in the removal of the 4th exon, causing a frame-shift mutation. Two rounds of gene targeting were performed to eliminate both alleles in this diploid cell line. Red arrows depict PCR primers to monitor gene status. Red triangles represent LoxP sites. B) PCR confirming the conversion of one wild-type (WT) allele to a null allele

in a *PARP1*$^{+/-}$ cell and the conversion of both wild-type alleles to null in *PARP1*$^{-/-}$ null cells. C) Western blot confirmation of the loss of PARP1 expression and confirmation of complementation of the null cells with PARP1 protein. D) A PARP1 activity assay demonstrates that the wild type and indicated complemented cells exhibited WT-levels of parylation, while the null cells lacked such activity. E) Growth curve depicting that the absence of PARP1 results in a slow growth phenotype.

Figure 3.2

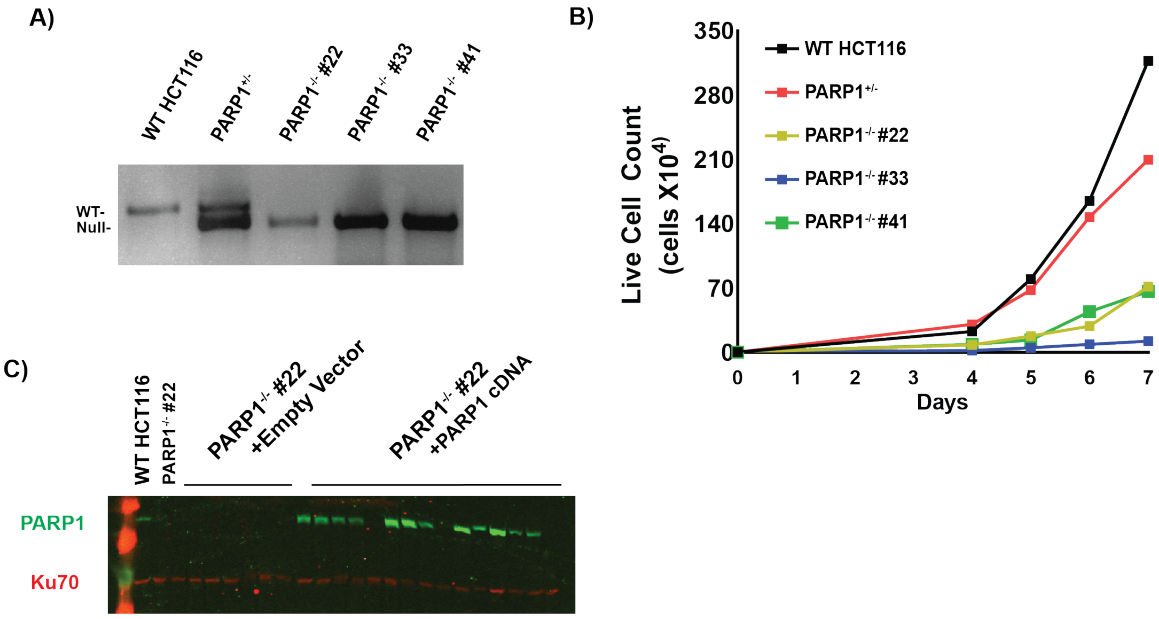


Figure 3.2. Characterization of the 3 independent *PARP*1-null cell lines.

A) A PCR analysis is shown using primers that flanked exon 4. The wild type (WT) allele PCR product migrates slightly slower than the PCR product corresponding to the null allele. The genotypes and clone designations are shown above the gel. B) PARP1-null cell lines show a significant proliferation defect. Cells from the indicated cell lines were seeded into dishes and then the total cell number was monitored by trypan blue staining at days 4 to 7 post plating. C) A Western blot analysis of the *PARP1*-null subclone #22 cell line. PARP1 protein is shown in green and as a loading control, the Ku70 protein is shown in red. The red smears on the far left in the unmarked lane are protein standard markers. PARP1-null clones expressing only the empty vector contained no detectable PARP1 protein whereas the complemented clones expressed various levels of PARP1 protein.

Figure 3.3

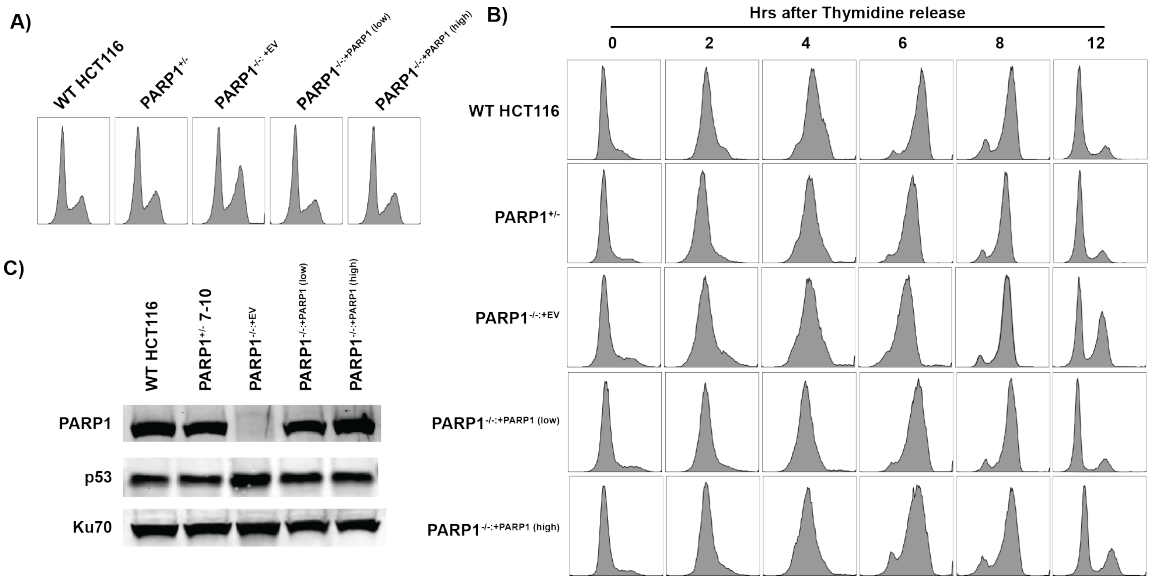


Figure 3.3. Parp1-null cells exhibit a G2-growth arrest.

A) The DNA content of asynchronously growing cells exhibited a modest, constitutive G_2 cell cycle accumulation. B) A time course study of thymidine-block synchronized cells. After release, all cells appeared to progress through S-phase at approximately the same rate, but many *PARP1*-null cells did not progress through mitosis, but rather exhibited a G_2/M cell cycle accumulation. C) Western blot evidence for increased p53 expression in PARP1$^{-/-}$ cells. Ku70 was used as a loading control.

Figure 3.4

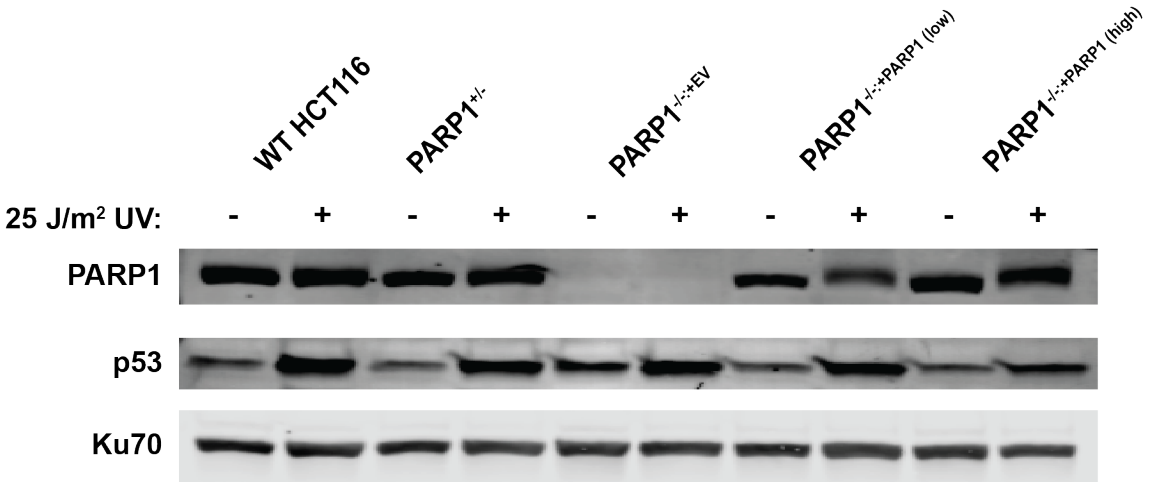


Figure 3.4. Levels of p53 protein in the *PARP*1-null cells.

The indicated cell lines were exposed to +/- 25 J/m^2 UV radiation, and whole cell lysates were collected 24 hr later. Western blot analyses were subsequently performed for PARP1, p53, and, as a loading control, Ku70. The *PARP1*$^{-/-}$ cells exhibited elevated levels of spontaneous p53 signaling. However, all the cell lines were capable of additional signaling via increased p53 expression following UV irradiation.

Figure 3.5

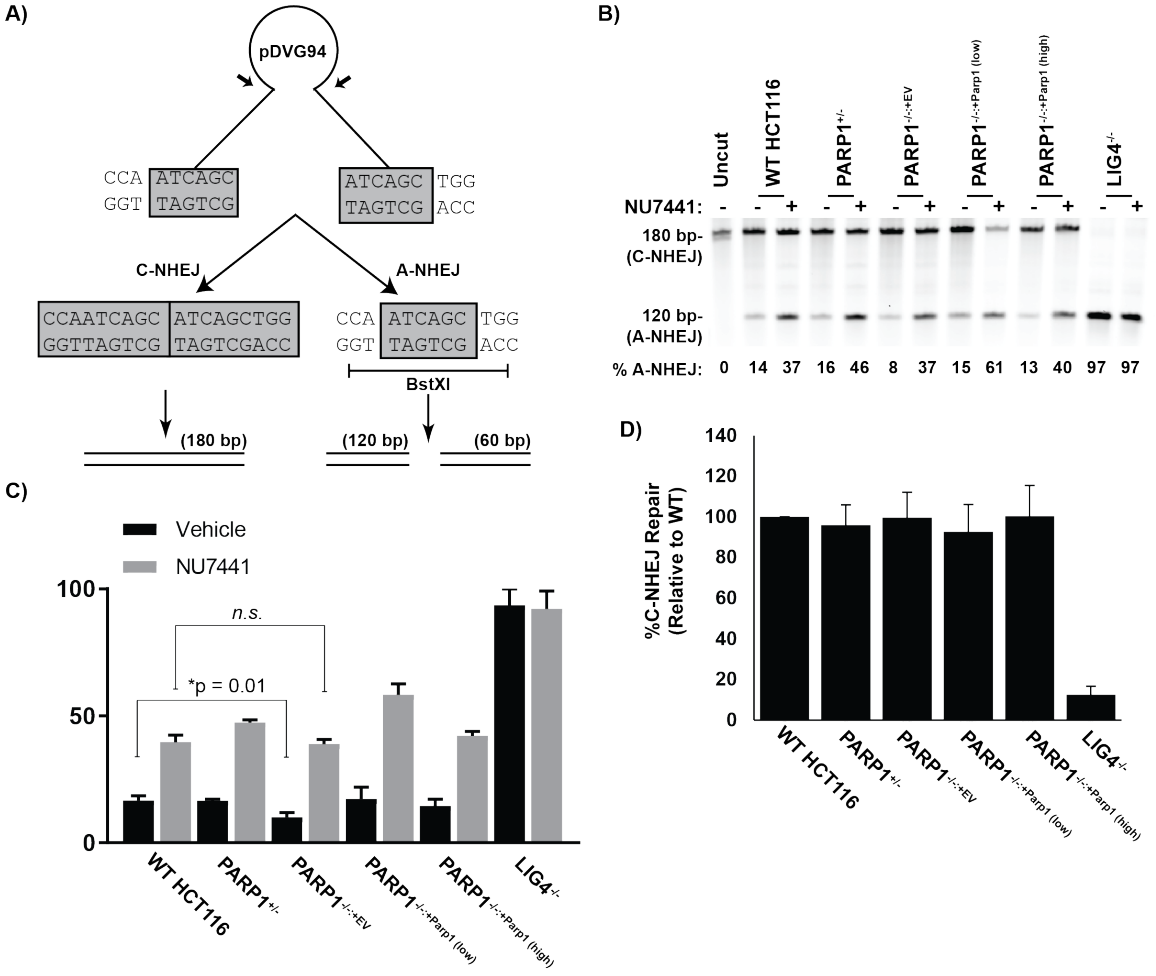


Figure 3.5. The impact of the absence of PARP1 on A- or C-NHEJ.

A) Schematic of the pDVG94 plasmid. B) The indicated cell lines were treated with 1 μM of the DNA-PKcs inhibitor, NU7441, for 4 hr, and then transfected with linearized pDVG94. Cells were allowed 24 hr to repair the linearized template (still in the presence or absence of inhibitor), then plasmids were extracted, and the region spanning the cut site was amplified by PCR, followed by digestion with *BstXI*. C) Quantitation of three experiments similar to panel B. *PARP1*$^{-/-}$ cells exhibited a significant (p = 0.01) ~2-fold reduction in baseline A-NHEJ activity but were not statistically different from wild type cells under

induced conditions. D) The indicated cell lines were transfected with linearized pGEM-Ad2-EGFP plasmid and subjected to flow cytometry analysis. Only DNA *LIG4*$^{-/-}$ cells exhibited a significant defect in DNA repair.

Figure 3.6

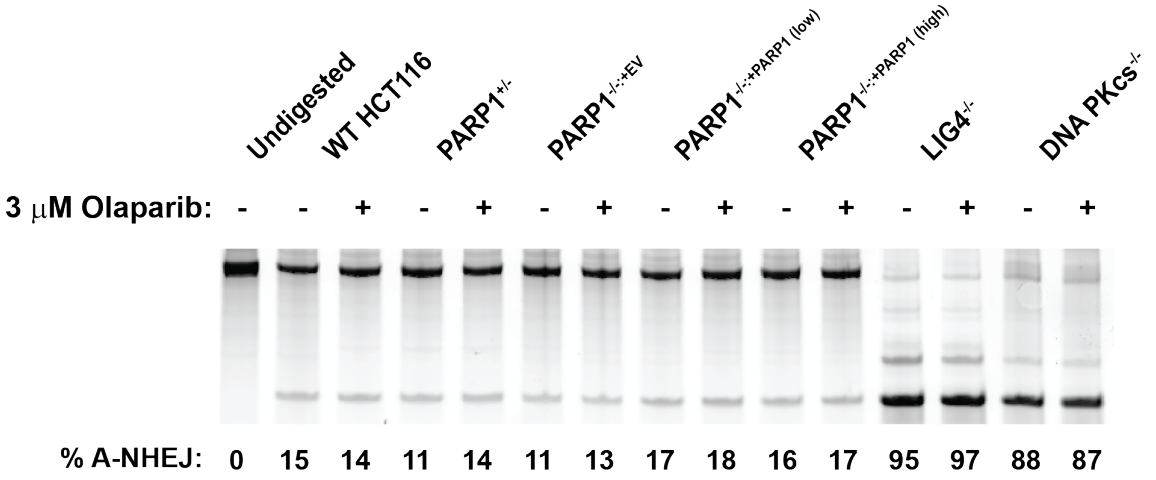
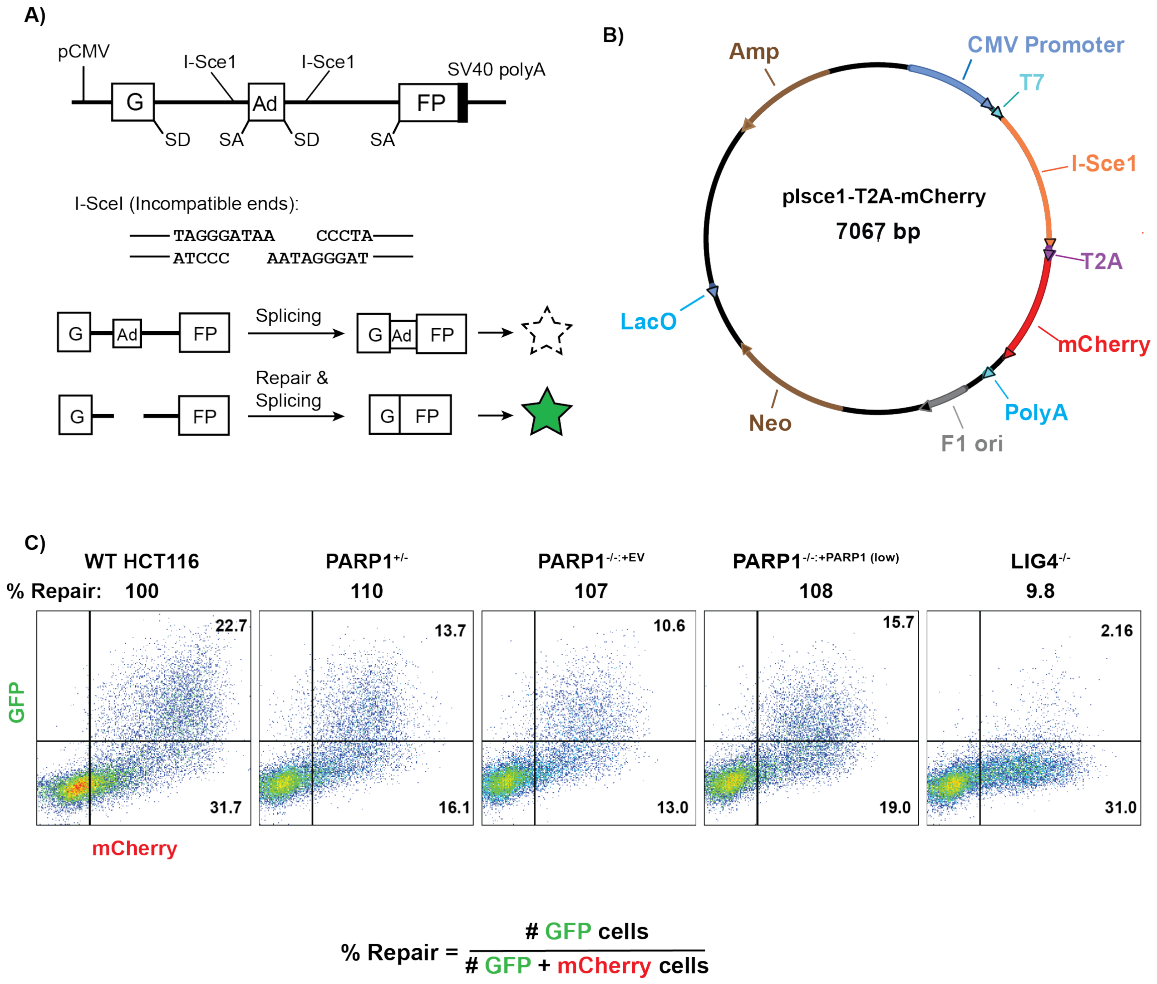


Figure 3.6. PARP1 inhibition does not affect A-NHEJ activity.

The indicated cell lines were either left untreated (-) or treated (+) with 3 μM of olaparib, a PARP1 inhibitor, for 4 hr, and then transfected with a linearized pDVG94 plasmid. The indicated cell lines were allowed 24 hr to repair the linearized template (still in the presence or absence of inhibitor), and then plasmids were extracted, and the region spanning the cut-site was amplified by PCR, digested with the *BstXI* restriction enzyme and then analyzed by gel electrophoresis. The % of A-NHEJ, as described in the legend to Fig. 3B, was determined and is shown below the gel. *N.B.* The sensitivity of HDR-defective (*FANCA*-null) cells to the same concentration of olaparib was used as a positive control to demonstrate that the olaparib used in this experiment was active (data not shown).

Figure 3.7



$$\% \text{ Repair} = \frac{\#\ \text{GFP cells}}{\#\ \text{GFP} + \text{mCherry cells}}$$

Figure 3.7. The absence of *PARP1* does not affect C-NHEJ activity.

A) Schematic of the pGEM-Ad2-EGFP NHEJ reporter. The rectangular boxes correspond to the indicated protein open reading frames. SD and SA indicate splice donor and acceptor sequences. This reporter must be linearized and repaired within the indicated intronic regions (resulting in the lack of retention of the Ad sequences), to enable GFP expression. B) A schematic of the mCherry vector used in this assay to control for transfection efficiency. C) A representative set of FACS images for the indicated cell lines. The numbers within the panels represent the fraction of cells that were GFP or RFP

positive, respectively. The formula for determining the percent repair is shown below the panels. Only the *LIG4*-null cells were deficient in this assay (see Fig. 3.5d).

Figure 3.8

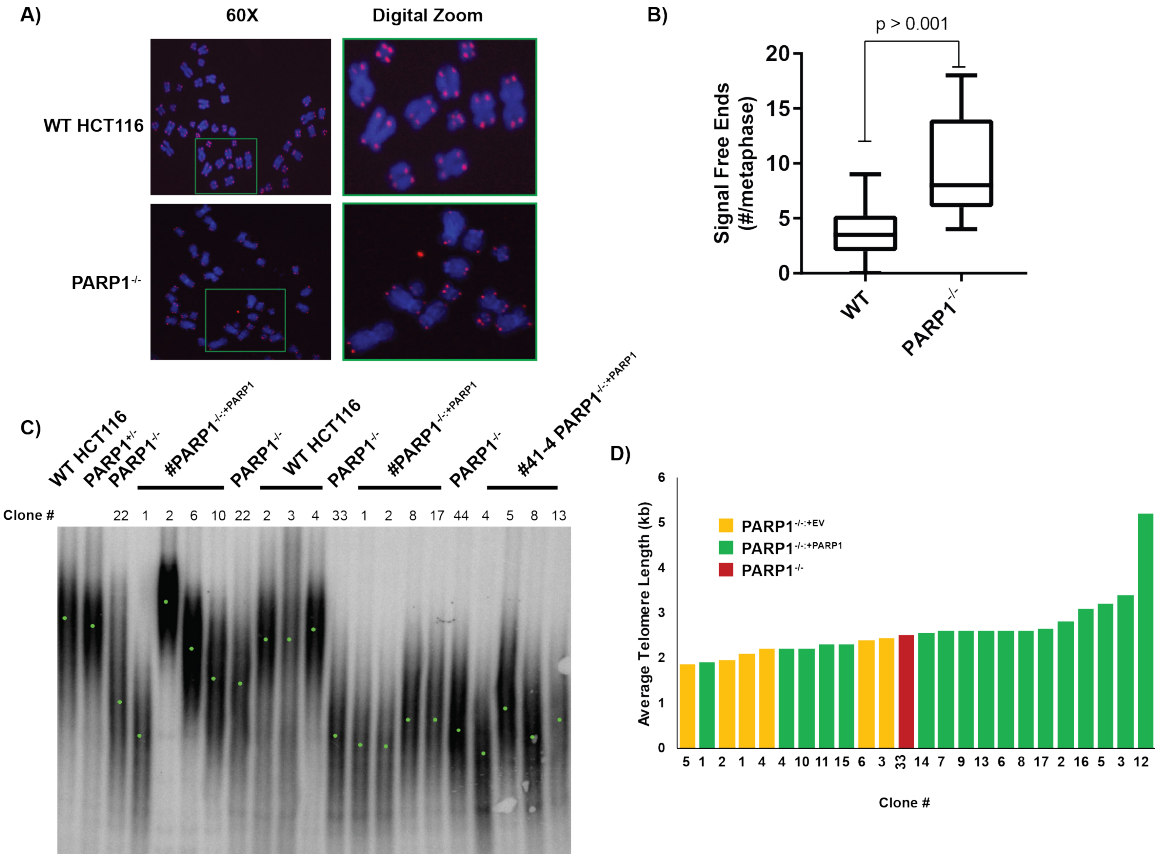


Figure 3.8. PARP1-null cells exhibit telomere dysfunction.

A) PARP1-null cells have an increased frequency of signal free ends, which is quantified in (B). Metaphase spreads were prepared from the indicated cell lines and then stain with a telomeric PNA probe (red spots) and then counter-stained with DAPI (blue). C) A TRF analysis of the the telomere length of the indicated cell lines. For many of the cell lines, independent subclones were isolated and these are indicated by the clone #. Genomic DNA from the indicated cell lines was prepared, digested to completion with frequent cutting restriction enzymes and the residual DNA was electrophoresed onto an agarose gel and then transferred to nitrocellulose. The blot was subsequently hybridized with a

radioactive telomeric probe. Since telomere length is variable from chromosome end to chromosome end and from cell to cell, a smear results. The mid-point of the telomeric smear is indicated with a green point. D) PARP1$^{-/-}$ cells have short telomeres, which can be rescued by complementation. In order to evaluate the clonal effect of telomere length, we derived empty vector containing (yellow rectangles), and complemented clones (green rectangles), from a given parental null clone (red rectangle) and then determined their telomere length by TRF analysis as shown in (C). The average telomere length by as determined by densitometry is shown.

Figure 3.9

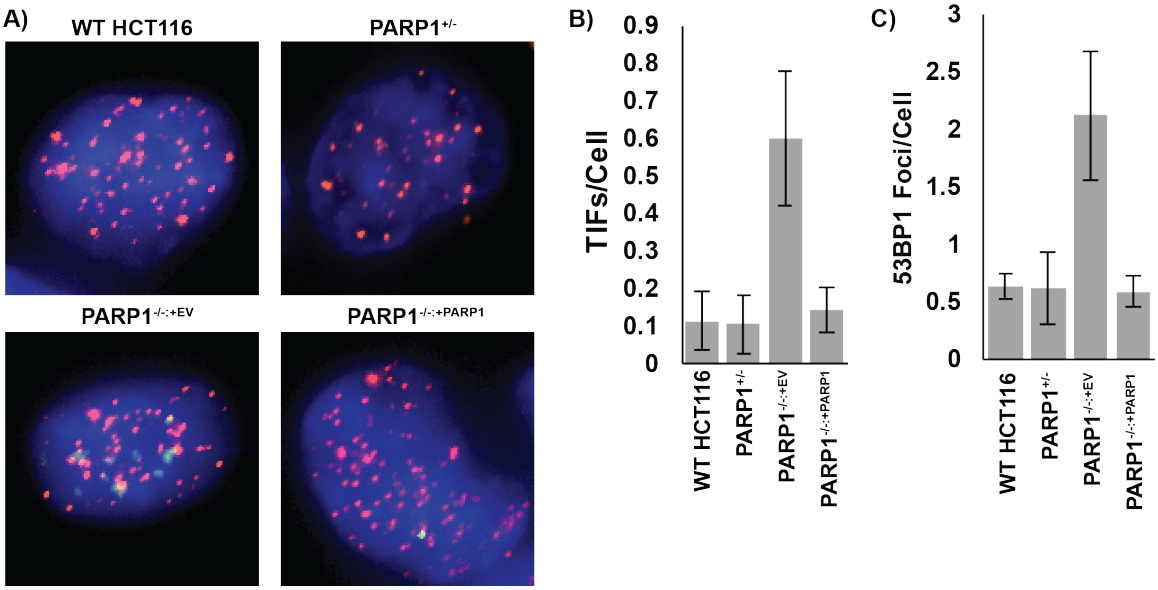


Figure 3.9. Spontaneous DNA damage foci in *PARP1*$^{-/-}$ cells co-localize with telomeres.

A) The indicated cells were immunostained for 53BP1 (green), fixed, and then subsequently probed for telomeric DNA with a PNA FISH probe (red) and for total DNA with DAPI (blue). PARP1$^{-/-}$ cells had an elevated level of 53BP1 foci, which tended to colocalize with telomeric DNA. B) The number of telomeric ends and overlapping 53BP1 foci (TIFs) on a per cell basis from images similar to panel (A) were averaged and graphed +/-1 standard deviation. *PARP1*$^{-/-}$ cells had a statistically significant increase in the frequency of TIFs compared to the control cell lines. C) Quantification of just 53BP1 foci/cell on a per cell basis from images similar to panel (A) were averaged and graphed +/-1 standard deviation. *PARP1*$^{-/-}$ cells had an increased level of endogenous 53BP1 foci, which was indicative of DNA damage.

Figure 3.10

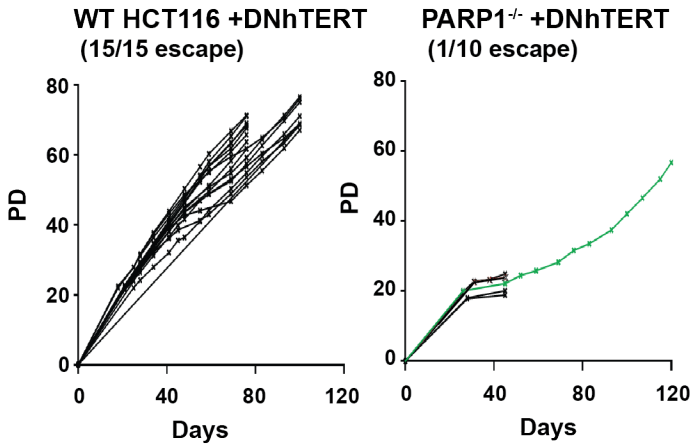


Figure 3.10. PARP1-/- cells are severely compromised in surviving telomeric stress. Growth curves plotting population doublings (PD) and days in culture. Cells were transduced with lentiviruses encoding DNhTERT, which is causative of telomere erosion over time. Each line represents an independent subclone. For some of the wild type cones the experiment was intentionally terminated after 80 days and for other only after 120 days. For PARP1-null cells, only a single subclone (green line) survived beyond 40 days.

Chapter 4

ATRX represses alternative lengthening of telomeres

[Published in: Christine E Napier[1,2], Lily I Huschtscha[1,2], Adam Harvey[3], Kylie Bower[1], Jane R Noble[1], Eric A Hendrickson[3] and Roger R Reddel[1,2] (2015). Oncotarget **6**(18): 16543-16558.

[1]Sydney Medical School, University of Sydney, NSW 2006, Australia

[2]Cancer Research Unit, Children's Medical Research Institute, Westmead, NSW 2145 Australia

[3]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota Medical School, Minneapolis, MN 55455 USA

Author Contributions:

- Harvey, A. was responsible for the creation of the *ATRX*$^{Neo/o}$ HCT116 cells, and the demonstration that they did not exhibit any ALT phenotypes. This work is presented in Figure 4.1 and Table 4.2. The description of this work in the Results and Methods sections was written by Harvey, A.
- The remainder of the work was either directly produced and conceived by Napier, C. or directed by her.
- Hendrickson, E. A. and Reddel, R. provided assistance with experimental design.

N.B. This work also served as the foundation for the work presented in **Chapter 5**. Thus, it's inclusion here was critical to fully representing the full scope of my role in demonstrating the role of ATRX in the suppression of ALT immortalization.

Synopsis

The unlimited proliferation of cancer cells requires a mechanism to prevent telomere shortening. ALT is a homologous recombination-mediated mechanism of telomere elongation used in tumors, including osteosarcomas, soft tissue sarcoma subtypes, and glial brain tumors. Mutations in the *ATRX*/*DAXX* chromatin remodeling complex have been reported in tumors and cell lines that use the ALT mechanism, suggesting that *ATRX* may be an ALT repressor. We show here that knockout or knockdown of *ATRX* in mortal cells or immortal telomerase-positive cells is insufficient to activate ALT. Notably, however, in SV40-transformed mortal fibroblasts ATRX loss results in either a significant increase in the proportion of cell lines activating ALT (instead of telomerase) or in a significant decrease in the time prior to ALT activation. These data indicate that loss of ATRX function cooperates with one or more as-yet unidentified genetic or epigenetic alterations to activate ALT. Moreover, transient ATRX expression in ALT positive/ATRX-negative cells represses ALT activity. These data provide the first direct, functional evidence that ATRX represses ALT.

Introduction

Telomeres are repetitive DNA structures located at the ends of chromosomes that shorten with each cell division (Harley *et al.* 1990). When telomeres become sufficiently shortened, the cell will enter a permanent proliferation arrest termed senescence. In order to become immortalized, cancer cells require a mechanism of telomere length maintenance. There are two mechanisms currently known to maintain telomere length in human cells: the reverse transcriptase enzyme telomerase and the recombination-mediated DNA synthesis mechanism called ALT (Greider and Blackburn 1985, Bryan et al. 1995). Telomerase activation can occur through numerous means, including loss of repressors, transcriptional upregulation or amplification of the genes encoding a telomerase subunit, TERT and/or TERC, and mutations in the TERT promoter (Reddel 2014). Somatic cell hybridization analyses showed that ALT activation occurs through loss of one or more repressor molecules that are present in normal somatic cells and in telomerase positive cells (Bryan et al. 1995, Perrem *et al.* 1999). More recently, it was found that inactivating mutations in one or other of the genes encoding the ATRX/DAXX chromatin remodeling complex, most commonly ATRX, are very common in ALT positive tumors and cell lines (Heaphy et al. 2011, Cheung *et al.* 2012, Lovejoy et al. 2012, Schwartzentruber *et al.* 2012, Chen *et al.* 2014). ATRX loss was also highly correlated with ALT in a panel of 19 ALT/telomerase cell line hybrids (Bower et al. 2012). ATRX has been proposed to have numerous diverse functions, including chromatin remodeling, viral resistance, and fidelity of chromatin separation during cell division, as well as binding to tandem DNA repeats (Xue *et al.* 2003, Ritchie *et al.* 2008, Law *et al.* 2010, Lukashchuk

and Everett 2010). Nonetheless, these data suggest the hypothesis that ATRX may be an ALT repressor.

In human cell culture models of immortalization, the first events required for this process usually involve inactivation of the p53 and pRB/p16INK4a tumor suppressor pathways by, for example, expression of one or more oncoproteins such as the SV40 large T antigen, which allows the cells to continue dividing beyond the point at which control cultures become senescent. After a finite number of additional cell divisions, however, the cells cease proliferation and the culture enters crisis, until a rare cell (approximately 1 in 10^7 fibroblasts and 1 in 10^5 epithelial cells) escapes crisis through spontaneous activation of a TEM, and the cell culture becomes immortal (Girardi *et al.* 1966, Huschtscha and Holliday 1983, Ide *et al.* 1984, Counter *et al.* 1992). In fibroblast cultures, the probability that ALT will be activated is approximately the same as for activation of telomerase, whereas epithelial cells are much more likely to activate telomerase; this reflects the situation in human cancers where carcinomas are usually telomerase-positive and ALT is most prevalent in cancers of mesenchymal origin (Colgin and Reddel 1999). It has long been assumed that activation of a TEM and escape from crisis involves spontaneous genetic or epigenetic events, that different events are required for activation of ALT and telomerase and, based on the frequency of their occurrence, that activation of each TEM requires one or more "hits".

In this study, we determined whether loss of ATRX is one of the genetic events involved in activating ALT, comparing epithelial cells with fibroblasts. To this end, we genetically disrupted ATRX in immortal telomerase positive epithelial cells to determine whether loss of ATRX was sufficient to activate ALT. As these cells did not activate the ALT mechanism, we went on to knock down ATRX in SV40-transformed pre-crisis cells.

In cells of epithelial origin, a reduction in ATRX also did not lead to ALT activation. However, in pre-crisis SV40-transformed fibroblasts derived from two different sources, the knockdown of ATRX led either to an increased frequency of ALT activation or a decrease in the time of crisis prior to immortalization via the ALT mechanism. Just as importantly, the transient expression of exogenous *ATRX* in three independent ALT-positive, *ATRX*-negative cell lines led to a reduction in C-circles and APBs, two markers of ALT activity. These data provide the first functional evidence that ATRX acts as a repressor of the ALT mechanism in cells of mesenchymal origin.

Results

ATRX gene knockout does not activate ALT in a telomerase-positive cell line

Previous findings in tumor cells and cell lines have shown a correlation between ALT and absence of ATRX at the protein or gene level. To determine whether knockout of ATRX in the telomerase-positive epithelial cell line HCT116 activated the ALT mechanism, we used two targeting strategies: CRISPR/CRISPR-associated systems (Cas)9 and recombinant adeno-associated virus (rAAV). Following single cell cloning of HCT116 cells co-transfected with CRISPR/Cas9 (Chen *et al.* 2013), correctly targeted clones were identified by restriction enzyme analysis of a PCR product that encompassed the CRISPR target (Figure 4.1A). HCT116 clones with a disrupted *ATRX* were resistant to digestion with the *Sml*I restriction enzyme due to removal of the *Sml*I recognition site, which was verified by Sanger sequencing. As a second method to knock out *ATRX*, an rAAV gene targeting vector was constructed to replace exon 5 of *ATRX*, and four clones resistant to G418 were analyzed for correct integration of the construct using primers that flank the proposed integration site (Figure 4.1B). Absence of ATRX protein expression in knockout clones constructed using either the CRISPR/Cas9 or rAAV method was confirmed by Western blot (Figure 4.1C).

We analyzed the *ATRX* knockout HCT116 cells to determine whether the ALT mechanism was activated. Similar to wild-type HCT116 cells, *ATRX* knockout HCT116 cells were negative for C-circles, partially single-stranded circles of C-rich telomeric DNA that are highly associated with the ALT mechanism (Henson et al. 2009) (Figure 1D). We also found all *ATRX* knockout HCT116 clones expressed telomerase activity as determined by the TRAP assay (Figure 4.1E). Telomere length was assessed by

Southern blot 29 population doublings (PDs) after cloning and did not show the elongated and heterogeneous telomere length profile that is characteristic of cells that utilize ALT (Figure 4.1F). These data demonstrate that the knockout of ATRX in telomerase-positive carcinoma cells is not sufficient to activate the ALT mechanism.

ATRX depletion in epithelial cells does not promote ALT activation

Epithelial-derived cell lines and tumors preferentially activate telomerase (Colgin and Reddel 1999, Henson et al. 2009), and cells that are telomerase-positive generally express ATRX (Bower et al. 2012, Lovejoy et al. 2012). We therefore determined whether loss of *ATRX* promoted ALT activation in SV40-transformed mammary epithelial cells. Two epithelial cell strains derived from individual SV40-transformation events, Bre80 T5 and Bre80 T8, were transduced with an empty vector (vector), scrambled control shRNA (sc), green fluorescent protein (GFP), or an shRNA targeting ATRX or DAXX. One of five control cultures (Bre80 T5 GFP; Table 1), and six of ten shATRX or shDAXX cultures emerged from crisis. Each immortal line was assessed for ATRX and DAXX expression (Figure 4.2A). Bre80 T5 GFP cells expressed ATRX and DAXX protein. The four shATRX and two shDAXX immortal lines lacked ATRX and DAXX expression, respectively.

All immortal cultures exhibited a period of crisis that ranged from 64 to 166 days (Figure 4.2B). All immortal cultures were positive by the TRAP assay for telomerase activity (Figure 4.2C) and negative for the presence of C-circles (Figure 4.2D), together indicating that all seven cultures activated telomerase. These results demonstrate that ATRX depletion in epithelial cells does not promote ALT activation. Furthermore, these data, together with the HCT116 *ATRX* knockout data, indicate that neither telomerase activity nor activation of telomerase is dependent upon the presence of ATRX.

ATRX knockdown in fibroblasts increases the proportion of cells activating ALT

Since we did not obtain any immortal epithelial cell lines that had activated the ALT mechanism, we next examined the effect of ATRX depletion in fibroblasts. The prevalence of ALT in tumors is skewed towards those of mesenchymal origin (Ulaner *et al.* 2003, Henson and Reddel 2010, Heaphy *et al.* 2011). We therefore determined whether fibroblasts derived from the same breast sample as Bre80 epithelial cells, with a similar genetic background, had a greater propensity toward ALT activation when ATRX expression was depleted. Two independent SV40-transformed, mortal breast fibroblast cultures, Fre80-3T and Fre80-4Tii, were transduced with either a vector encoding a scrambled shRNA sequence or an shATRX vector. Transduced cells and untransduced (parental) controls were maintained in culture until cultures either escaped from crisis and became immortal, or failed to escape from crisis and became non-viable. ATRX and DAXX proteins were both expressed by each control (scrambled shRNA or parental) culture, and ATRX was efficiently suppressed in each shATRX culture at a pre-crisis time point (Figure 4.3A). Six out of 13 control cell strains became immortalized, whereas 12 of 16 shATRXtransduced cultures became immortal (Table 4.1). Expression of ATRX and DAXX proteins in the immortalized cultures (Figure 4.3B) remained similar to that of pre-crisis cells. All immortal cultures exhibited a period of crisis, and the mean length of crisis between control and shATRX cultures was not significantly different (50 versus 52 days; Table 4.1). The TEM of each immortal culture was determined using the TRAP and C-circle assays. Of six control cultures that became immortal, five cultures exhibited telomerase activity by the TRAP assay, while the remaining culture was negative for telomerase activity (Figure 4.3C). Only two of 12 immortal shATRX cell lines showed telomerase activity. The results of the C-circle assay correlated inversely with those of

the TRAP assay: each TRAP positive sample was negative for C-circles, and *vice versa* (Figure 4.3D). The results demonstrate that the induced loss of ATRX significantly promotes ALT activation, as 10 of 12 shATRX-transduced cultures activated the ALT mechanism, while only one of six control cultures was ALT-positive (p = 0.01, Fisher's exact test). These data provide the first functional evidence that, in fibroblasts, ATRX loss facilitates ALT activation.

ATRX knockdown decreases the time required for occurrence of immortalization

We then depleted ATRX in two clonal SV40-transformed pre-crisis fibroblast strains from a different source. In addition, we also knocked down DAXX, as both proteins act together as chromatin remodelers and one or both is mutated in pancreatic neuroendocrine tumors with an ALT-like phenotype (Lewis *et al.* 2010, Heaphy et al. 2011). ATRX and DAXX proteins were expressed by both pre-crisis strains (JFCF6/T.1/P and JFCF-6/T.5K) (Figure 4.4A, lanes labeled parental and mortal). shATRX and shDAXX lentivirus were used to efficiently knock down ATRX or DAXX in both fibroblast cell strains (Figure 4.4A, shATRX and shDAXX mortal samples). Transduction of the empty vector (vector) or scrambled shRNA control (sc) did not affect endogenous ATRX or DAXX expression. Each mortal culture was passaged through a period of crisis until it became immortal. Growth curves were plotted for each cell line to examine whether there was a change in the length of crisis in shATRX or shDAXX cultures compared to controls (Figure 4.4B). Six out of eight control cultures showed a distinct period of crisis, ranging from 13 to 78 days (Table 4.1). Compared to immortal control cultures, shATRX- or shDAXX-transduced cell lines became immortalized after a significantly reduced length of time in crisis (range: 0 to 28 days; $p < 0.05$, Mann Whitney test).

Spontaneous loss of ATRX expression is also associated with the activation of ALT

ATRX and DAXX protein expression was analyzed in each immortal JFCF-6 cell line (Figure 4A, immortal lanes). ATRX expression was spontaneously lost in 7 of 8 immortal control cultures, as well as in one immortal shDAXX culture. In contrast, spontaneous loss of DAXX was not observed in any immortal culture. ATRX knockdown was maintained in all shATRX-transduced cultures after they became immortalized. Similarly, substantial knockdown of DAXX was maintained after immortalization of both shDAXX-transduced cultures. We sequenced all 35 exons of ATRX to determine whether ATRX protein loss was due to mutation, and identified a premature stop codon in two cell lines that spontaneously lost ATRX expression (ATRX exon 9 of the JFCF-6/T.5K-vector cell line and ATRX exon 10 of the JFCF-6/T.5K-shDAXX culture). The ATRX sequence was wildtype in the remaining six immortal cultures that spontaneously lost ATRX expression, indicating that in these cells ATRX protein is not expressed for reasons other than changes in the coding sequence.

We examined the temporal correlation between spontaneous loss of ATRX expression and crisis in three JFCF-6/T.1/P lines, two of which (unmodified parental and vector-transduced) spontaneously lost, and one of which (sc1) maintained ATRX protein expression after immortalization (Figure 4.5). In both JFCF-6/T.1/P-parental and -vector lines, spontaneous loss of ATRX occurred early during culture crisis. In contrast, the JFCF-6/T.1/P-sc1 culture maintained ATRX expression through crisis. These data demonstrate that spontaneous loss of ATRX can be an early event in the process of cellular immortalization.

The TEM that was activated in each immortal JFCF-6/T.1/P- and JFCF-6/T.5K-derived culture was assessed. Every culture was negative for telomerase activity, both before and

after immortalization, as demonstrated by the TRAP assay (Figure 4.6A). All cultures were negative for C-circles prior to immortalization and all post-crisis cultures were C-circle positive (Figure 4.6B). Telomere length was assessed by Southern blot in each cell line at a minimum of two time points, before and after immortalization, and in each case the immortalized cultures exhibited the heterogeneous telomere length pattern characteristic of ALT, contrasting with the more homogeneous telomere lengths in the pre-crisis cells (Figure 4.6C). Thus, all immortal JFCF-6-derived fibroblast cell lines activated the ALT mechanism: eight control and 10 shATRX/shDAXX cultures. Spontaneous loss of ATRX was observed in 7 of 8 control cultures and in one shDAXX cell line. These data are clear confirmation that ATRX loss facilitates ALT activation in fibroblasts.

ATRX expression represses the ALT phenotype

Given that the loss of ATRX promoted activation of the ALT mechanism in fibroblasts, we wanted to determine whether restoration of ATRX expression would repress the ALT phenotype. Exogenous ATRX was expressed in three ALT cell lines of mesenchymal origin that lack ATRX expression (GM847, JFCF-6/T.5K-sc1 and U-2 OS). ATRX protein expression was examined at 2, 4, 6 and 8 days following transfection with an empty vector (EV) or an ATRX expression construct. Maximal ATRX protein expression was observed 2 days following transfection, and decreased rapidly, returning to undetectable levels by day 6 in all cell lines examined (Figure 4.7A). Exogenous expression of ATRX did not affect DAXX expression in any cell line. Furthermore, ATRX transfection did not affect the growth rate or cell cycle kinetics in any of the three cell lines tested (data not shown).

As C-circle levels are indicative of ALT and rapidly respond to perturbations in ALT activity (Henson et al. 2009), we determined whether C-circle levels changed in response

to ATRX transfection. The level of C-circles was significantly reduced following ATRX expression in each of the three cell lines examined compared to untreated, FuGENE-treated or EV-transfected cells (Figure 4.7B).

APBs are highly correlative with the ALT mechanism (Yeager et al. 1999), and previous studies have demonstrated inhibition of ALT activity results in a decrease in APBs (Jiang *et al.* 2005). Therefore, as a further indication that ATRX expression affects the ALT phenotype, APBs were quantified in ALT cells transiently expressing ATRX. The percentage of ATRX-positive nuclei that also contained APBs was significantly reduced compared to the level of APB-positive nuclei in EV-transfected cultures at days 2 and 4 in all cell lines examined (Figure 4.7C). By day 6, there were insufficient ATRX-positive nuclei to count ATRX-positive/APB-positive nuclei. These results provide further direct evidence that ATRX represses the ALT mechanism.

As an additional control, GM847 cells were transfected either with an EV or a plasmid encoding GFP. The percentage of GM847 nuclei that were positive for both GFP and APBs (80, 81 and 80% at days 2, 4 and 6, respectively) was not significantly different to the percentage of APB-positive nuclei in an EV-transfected culture (75, 87 and 81% at days 2, 4 and 6, respectively). These data confirm that the percentage of nuclei containing APBs was not affected by expression of an irrelevant exogenous protein. We also examined whether transient ATRX expression affected mean telomere length or the overall telomere length heterogeneity. Neither the mean telomere length nor the telomere length heterogeneity of GM847, JFCF-6/T.5K-sc1 or U-2 OS was affected by the transient expression of ATRX (Figure 4.7D). Therefore rapid changes in telomere length did not occur in response to transient ATRX expression in ALT cells, and long-term ATRX expression may be required for changes in telomere length to occur.

Discussion

Mutations in members of the ATRX/DAXX chromatin remodeling complex have been implicated in the ALT mechanism, in both ALT positive cell lines and tumor samples (Henson and Reddel 2010, Heaphy et al. 2011, Bower et al. 2012, Lovejoy et al. 2012). Although a previous study was unable to demonstrate immortalization of SV40-transformed BJ fibroblasts following ATRX knockdown (Lovejoy et al. 2012), we found that depletion of ATRX facilitates immortalization of fibroblasts, increasing the proportion of cells that activate ALT compared to telomerase activation (Fre80 cultures) and decreasing the time in crisis prior to immortalization (JFCF-6 cultures). There was no evidence that the effect of ATRX depletion was different in these two sets of cultures: there was an insufficient number of Fre80 cultures which spontaneously activated ALT in the absence of ATRX knockdown to see a significant effect on length of crisis and, conversely, all of the spontaneously immortalized JFCF-6 cultures in this study activated ALT, making it impossible to observe an increased proportion of ALT-positive cultures following ATRX knockdown. The data are consistent with the hypothesis that 1) activation of ALT and telomerase in the context of inactivated p53 and pRb/p16INK4a tumor suppressor pathways each requires at least two genetic or epigenetic events, 2) that the events are different for the two TEMs, and 3) that one of the events for activating ALT is the loss of ATRX function. In pre-crisis cells that have undergone none of the events required for activation of either ALT or telomerase, experimentally depleting ATRX would result in fewer additional events being required for activation of ALT, which would both increase the probability that ALT becomes activated and decrease the time required for immortalization to occur.

Spontaneous ATRX loss was observed in eight immortal JFCF-6 fibroblast cultures; the time course of ATRX loss was examined in two of these cultures, and in both cases this occurred early in crisis. This is also consistent with the hypothesis that ATRX loss-of-function is one of at least two genetic events required for escape from crisis via activation of ALT. It is not possible to conclude from this limited number of observations whether ATRX loss is usually the first event, possibly contributing to the probability that the additional event(s) will occur, or whether the order of the events is stochastic. The reason for spontaneous ATRX loss was determined to be a mutation that resulted in a premature stop codon in two immortal lines. The absence of an exonic mutation in the other lines is consistent with a previous report of 12 ALT positive cell lines (including two JFCF-6 lines) that had wild-type ATRX sequence but lacked or had abnormal ATRX protein expression (Lovejoy et al. 2012). Absence of ATRX protein expression in the presence of wild-type ATRX sequence may potentially be due to epigenetic changes, miRNA targeting or post-translational modifications. The presence of ATRX expression in a small number of ALT-positive cell lines suggests that there are other proteins whose loss can have the same functional outcome as loss of ATRX expression. The higher frequency of ALT activation in JFCF-6 compared to Fre80 cells may indicate that the respective pre-crisis cultures may have accumulated different pre-disposing genetic or epigenetic events. Nevertheless, the activation of ALT in all cell lines that spontaneously lost ATRX provides additional evidence for its role as an ALT suppressor.

Successful knockout of ATRX using CRISPR/Cas9 technology or rAAV in epithelial-derived HCT116 colon carcinoma cells did not affect telomerase expression nor activate ALT. This indicates that telomerase activity does not require the presence of functional ATRX. The lack of any ALT immortalized epithelial cell lines in the breast epithelial cell

experiments, in contrast to the results with breast stromal fibroblasts derived from the same individual, is consistent with the observation that ALT positive tumors are more frequently of mesenchymal, rather than epithelial, origin (Henson *et al.* 2002) and support the concept that there are inherent cell type differences, and possibly requirements, with regard to cellular immortalization (Lafferty-Whyte *et al.* 2009). The genetic events required to activate telomerase in cell types, including epithelia, where low levels of telomerase activity occur physiologically, may differ from the events required to activate telomerase in cell types such as fibroblasts where telomerase activity is normally undetectable, and presumably tightly repressed. If this is the case, then in epithelial cells where ATRX is experimentally depleted, the probability of the additional events required to activate telomerase occurring may still be greater than the probability of activating ALT.

Transient ATRX expression repressed the ALT mechanism in three separate ALT-positive/ATRX-negative cell lines. This is the first functional evidence that ATRX expression is able to repress the ALT mechanism. ALT repression clearly resulted from ATRX expression, as when ATRX was expressed, both C-circle levels and the proportion of APB positive nuclei decreased. Both C-circle and APB levels returned to control levels following the loss of ATRX protein overexpression. The mechanism by which ATRX represses the phenotypic ALT markers is unknown but may include facilitating replication and resolving G-quadruplex secondary structures (Law et al. 2010). An attractive alternative hypothesis is that the chromatin remodeling capabilities of ATRX are critical, as alterations in the heterochromatic state of telomeres and chromatin-associated proteins have been implicated in the ALT mechanism (Jiang *et al.* 2011, Conomos et al. 2014, Episkopou *et al.* 2014, O'Sullivan et al. 2014). Thus, silencing of the histone chaperone ASF1 induces the ALT phenotype (O'Sullivan et al. 2014) and the NuRD-ZNF827 complex

is present at ALT telomeres and causes histone hypoacetylation, a marker of closed chromatin (Conomos et al. 2014). Confusingly however, decondensation of telomeric chromatin, as determined by micrococcal nuclease digestion, has also been shown in ALT cells (Episkopou et al. 2014). Despite these somewhat contradictory results implying that ALT telomeres may be more heterochromatic or euchromatic depending upon the model system utilized, a consistent feature is that chromatin is somehow altered in all these cells and that this alteration seems likely causative, rather than simply an indirect effect. Further studies that determine the requirement for altered telomeric chromatin status in ALT, and how chromatin remodelers play an essential role in this process, are likely to yield important insights into the ALT mechanism.

Materials and Methods

Cell culture

JFCF-6/T.1/P and JFCF-6/T.5K cell strains were derived from independent SV40transformation events of mortal jejunal fibroblasts from a male cystic fibrosis patient. GM847DM cells (referred to as GM847) were derived from Lesch-Nyhan syndrome human fibroblasts immortalized with SV40 (Pereira-Smith and Smith 1988). HT1080, 293 and U-2 OS cells were obtained from the American Type Culture Collection (ATCC). Two separate SV40-transformed breast stromal fibroblast cell strains, Fre80-3T and Fre80-4Tii, were also used in the current studies (Kaul *et al.* 2011). All of these strains and lines were cultured in DMEM with 10% fetal bovine serum (FBS). Two individual SV40-transformed Bre80 cell strains, breast stromal epithelial cells derived from the same individual as the Fre80 fibroblasts, were cultured in MCDB-170 serum-free medium (Life Technologies). The HCT116 colon carcinoma cell line (ATCC) and its derivatives were cultured in McCoy's 5A medium supplemented with 10% FBS, penicillin and streptomycin (each 100 U/mL), and 5 mM L-glutamine. All cultures were grown in a humidified incubator at 37^{o}C with 5% CO_2. Crisis was defined as a time period during which the cell culture underwent less than 1 PD per 7 days. The cell lines were confirmed to be free of Mycoplasma species and the identity confirmed using short tandem repeat profiling by CellBank Australia (Sydney, Australia).

Lentiviral infection

Lentiviral constructs encoding a scrambled control shRNA in one of two vectors (pLKO.1 or pLKO.5), GFP (pLKO.1 vector), shATRX [constructs 13590 (pLKO.1 vector)

and 342811 (pLKO.5 vector)] or shDAXX [construct 3800 (pLKO.1 vector)], as well as the pLKO.1 empty vector were obtained from Sigma. Lentivirus was produced and target cells were infected with equivalent multiplicities of infection. Infected cells were continuously cultured in medium supplemented with 0.5 µg/mL puromycin. Three sublines were established from each JFCF-6/T.1/P- and JFCF-6/T.5KshATRX-1 mass culture prior to immortalization. The following cell strains/lines were transduced with pLKO.1-scrambled: JFCF-6/T.1/P-sc1, JFCF-6/T.5K-sc1, Bre80 T5 sc, Fre80-3T sc1-3 and Fre80-4Tii sc1-3; the remaining scrambled-shRNA control lines were transduced with a pLKO.5scrambled shRNA vector. Fre80-3T shATRX-1, Fre80-4Tii shATRX-4, Bre80 T5 shATRX-2 and Bre80 T8 shATRX-3, -4, -5 lines were transduced with shATRX construct 342811, while the remaining shATRX lines were transduced with shATRX construct 13590.

Transfection

An ATRX expression vector, pCMV6-Entry-ATRX, the empty vector pCMV6-Entry and a GFP-expressing vector, pCMV6-AC-GFP, were obtained from OriGene. Transfection with either vector was performed using FuGENE reagent (Promega). Cells were harvested at the indicated time points post-transfection and analyzed as described. The mean percentage (± SEM) of ATRX-positive nuclei at day 2 in GM847, JFCF-6/T.5K-sc1 and U-2 OS cells was 50 ± 11, 18 ± 2 and 26 ± 2%, respectively. At day 4, the mean ATRX-positive nuclei percentage (± SEM) in GM847, JFCF-6/T.5Ksc1 and U-2 OS cells was 39 ± 17, 10 ± 2 and 19 ± 4%, respectively.

CRISPR/Cas9 ATRX knockout

The ATRX cDNA was screened by ZiFiT to determine a CRISPR target in exon 9 of *ATRX* (Hwang *et al.* 2013). Two complementary oligos were designed by ZiFit, ATRXex9_1 and ATRXex9_2 (Table 4.2), and were subsequently annealed and ligated into a CRISPR RNA expression plasmid (MLM3636; Addgene). The resulting plasmid was co-transfected with a Cas9 nuclease expression plasmid (41815; Addgene) into wild-type HCT116. Cells were subcloned, and screened for correct targeting by interrogating the disruption of an *Sml*I restriction enzyme site that lies directly adjacent to the target sequence cut by the Cas9 endonuclease. Targeting PCR was performed using ATRXex9F and ATRXex9R (Table 4.2), and products were subjected to *Sml*I digestion. Sanger sequencing of the correctly modified clones was performed with ATRXex9SeqF to confirm that an early stop codon was inserted or a frame-shift mutation had occurred.

rAAV-mediated ATRX knockout

A knockout rAAV vector was constructed as described, with slight modifications (Kohli *et al.* 2004, Khan et al. 2011, Luo *et al.* 2014). The fifth exon of the *ATRX* was targeted, as it has been reported that disruption of this exon results in an ATRX knockout (Leung *et al.* 2013). Accordingly, the left and right homology arms were constructed by PCR with primer pairs, ATRX LarmF and ATRX LarmR, and with ATRX RarmF and ATRX RarmR, respectively (Table 4.2). These PCR products were combined with a neomycin resistance gene, and an rAAV vector backbone, and used in a Golden Gate ligation reaction to generate the rAAV ATRXKO. Following virus production and infection of wild-type HCT116 cells, neomycin resistant sub-clones were screened by PCR with ATRX TargF

and ATRX IresR primers (Table 4.2) to identify correctly targeted clones. Clones were continuously cultured in media containing 1 mg/mL G418.

Immunoblot analysis

Cell lysate preparation and protein detection was performed as described (Bower et al. 2012). Briefly, protein lysates were prepared from harvested cells, separated by electrophoresis using a Tris-acetate gel, transferred to PVDF membrane, and the membrane was probed for protein using the antibodies indicated in each figure. Antibodies used: ATRX (Sigma, HPA001906 or GeneTex, GTX629703), DAXX (Sigma, HPA008736), γ-tubulin (Sigma, T5192) and Ku70 (Santa Cruz, sc-9033).

APB detection

APB analysis was performed by cytocentrifuging cells onto SuperFrost Plus slides using a Shandon Cytospin 4. Cells were then fixed with 4% formaldehyde and permeabilized with 0.1% Triton X-100. Blocking and RNase A treatment were simultaneously performed by incubating slides with 0.1 mg/mL RNase A diluted in antibody dilution buffer (ABDIL; 20 mM Tris-HCl, pH 7.5, 0.2% fish gelatin, 2% BSA, 0.1% Triton X-100, 150 mM NaCl, 0.1% sodium azide). Cells were stained with antibodies specific for ATRX (Sigma, HPA001906), TRF2 (Merck Millipore, OP129) and PML (Santa Cruz, sc-9862) diluted in ABDIL for 2 h at 37°C in a humidified chamber. Following washes in PBS with 0.1% Tween-20 (PBST), cells were stained with appropriate AlexaFluor secondary antibodies for 30 min at 37°C in a humidified chamber. Subsequent to an additional set of PBST washes, nuclei were counterstained with DAPI and mounted using DABCO anti-fade mounting media. Images were obtained on an Imager.M1 microscope

(Zeiss) and analysis performed using ZEN software (Zeiss). Statistical analyses were performed using GraphPad Prism 5.0.

C-circle assay

The C-circle assay was performed as described (Henson et al. 2009). C-circle signal was quantified using ImageQuant software and background corrected by edge subtraction. Quantitation of the C-circle signal in the ATRX expression experiments was performed by normalizing the signal obtained for each sample to the amount of DNA used per reaction, followed by comparison to the signal obtained on the corresponding day post-transfection with the empty vector pCMV6-Entry.

Terminal restriction fragment (TRF) assay

Telomere length was assessed essentially as described (Bower et al. 2012). Equivalent amounts of genomic DNA were digested with *Hinf*I and *Rsa*I and separated using either standard or pulsed field gel electrophoresis. The resulting gel was hybridized either in-gel or after transfer to Hybond XL (GE Healthcare) membrane using a radioactive $(TTAGGG)_3$ oligonucleotide probe. Detection of radioactive signal was performed using a phosphor-imaging screen and the signal measured using a Typhoon Trio (GE Healthcare).

Telomere repeat amplification protocol (TRAP)

The TRAP assay was used to assess telomerase activity as described (Bower et al. 2012). Lysates were prepared using CHAPS buffer and equal amounts of lysate were used in each TRAP reaction. The TRAP products were separated on a PAGE gel and the gel was stained with SYBR Gold. Gels were scanned on a Typhoon Trio.

ATRX sequencing

ATRX was sequenced using described primers and conditions (Jiao *et al.* 2011). PCR products were purified and sent to the Australian Genome Research Facility for Sanger sequencing. Sequence analysis was performed using CodonCode Aligner.

Tables

Table 4.1. Pre- and post-crisis protein and telomere lengthening mechanism characterization of cell strains and cell lines.

Parental Cell Strain	Expression vector	Pre-crisis Protein		Days in Crisis	Post-crisis					TEM
					C-Circles	TRAP	TRF#	Protein		
		ATRX	DAXX					ATRX	DAXX	
Bre80 T5	GFP	nd	nd	166	-	+	-	+	+	TEL
	ev-1	nd	nd	136*	-					-
	ev-2	nd	nd	264*	-					-
	ev-3	nd	nd	264*	-					-
	sc	nd	nd	187*	-					-
	shATRX-1	nd	nd	109*	-					-
	shATRX-2	nd	nd	107	-	+	nd	-	+	TEL
	shATRX-3	nd	nd	64	-	+	nd	-	+	TEL
	shDAXX-1	nd	nd	98	-	+	nd	+	-	TEL
	shDAXX-2	nd	nd	77	-	+	nd	+	-	TEL
Bre80 T8	shATRX-1	nd	nd	92	-	+	-	-	+	TEL
	shATRX-2	nd	nd	97	-	+	-	-	+	TEL
	shATRX-3	nd	nd	102*						-
	shATRX-4	nd	nd	102*						-
	shATRX-5	nd	nd	102*						-
Fre80-3T	sc1	+	+	65	-	+		+	+	TEL
	sc2	+	+	23	-	+		+	+	TEL
	sc3	+	+	124*						-
	sc4	+	+	22	-	+		+	+	TEL
	sc5	+	+	99	-	+		+	+	TEL
	sc6	+	+	30	-	+		+	+	TEL
	shATRX-1	nd	nd	62	+	+		+	+	TEL
	shATRX-2	-	+	125*						
	shATRX-3	-	+	99	-	+	nd	-	+	TEL
	shATRX-4	-	+	158*						-
	shATRX-5	-	+	125*						-
	shATRX-6	-	+	158*						-
	shATRX-7	-	+	0	-	+	nd	low	+	TEL
Fre80-4Tii	none	+	+	60	+	-	ALT	+	+	ALT
	sc1	nd	nd	87*						-
	sc2	+	+	87*						-

Table 4.1 Pre- and post-crisis protein and telomere lengthening mechanism characterization of cell strains and cell lines. (Continued)

Parental Cell Strain	Expression vector	Pre-crisis Protein		Days in Crisis	Post-crisis					TEM
					C-Circles	TRAP	TRF#	Protein		
		ATRX	DAXX					ATRX	DAXX	
Fre80-4Tii	sc3	+	+	87*						-
	sc4	+	+	78*						-
	sc5	+	+	82*						-
	sc6	+	+	82*						-
	shATRX-1	nd	nd	31	+	-	+	-	+	ALT
	shATRX-2	nd	nd	93	+	-	nd	-	nd	ALT
	shATRX-3	nd	nd	112	+	-	+	-	+	ALT
	shATRX-4	-	+	14	+	-	+	-	+	ALT
	shATRX-5	-	+	27	+	-	nd	-	+	ALT
	shATRX-6	-	+	44	+	-	nd	-	+	ALT
	shATRX-7	-	+	57	+	-	nd	-	+	ALT
	shATRX-8	-	+	46	+	-	nd	-	+	ALT
	shATRX-9	-	+	42	+	-	nd	-	+	ALT
JFCF-6/T.1/P	None	+	+	44	+	-	+	-	+	ALT
	ev	nd	nd	13	+	-	+	-	+	ALT
	sc1	+	+	42	+	-	+	+	+	ALT
	sc2	+	+	78	+	-	+	-	+	ALT
	shATRX-1	-	+	0	+	-	+	-	+	ALT
	shATRX-2	-	+	0	+	-	+	-	+	ALT
	shATRX-3	-	+	0	+	-	+	-	+	ALT
	shATRX-4	-	+	0	+	-	+	-	+	ALT
	shDAXX	+	-	0	+	-	+	+	-	ALT
JFCF-6/T.5K	None	+	+	0	+	-	+	-	+	ALT
	ev	+	+	48	+	-	+	-	+	ALT
	sc1	+	+	46	+	-	+	-	+	ALT
	sc2	+	+	0	+	-	+	-	+	ALT
	shATRX-1	-	+	12	+	-	+	-	+	ALT
	shATRX-2	-	+	26	+	-	+	-	+	ALT
	shATRX-3	-	+	28	+	-	+	-	+	ALT
	shATRX-4	-	+	21	+	-	+	-	+	ALT
	shDAXX	+	-	0	+	-	+	-	-	ALT

TRF# Telomers had ALT like length
ev Empty Vector
sc Scrambled control

\+ Positive signal
\- Negative signal
nd Not determined

Table 4.2. List of PCR primers

Oligo Name	Sequence (5'-3')
ATRX LarmF	atacatagcggccgcaggtcctttgcctgatacac
ATRX LarmR	tattactagtgaccctggaggacactcttc
ATRX RarmF	ataactcgaggatgagaatgacagggaacttgg
ATRX RarmR	atattgcggccgcgtgaattgaagagagaggctgt
ATRX TargF	gaggtggatgttaagggcatg
ATRX IresR	gcttccagaggaactgcttcc
ATRX IntF	ctaaccagtgggaatgtctc
ATRX IntR	cagggaacaaggtttaataagcac
ATRXex9_1	acaccgtttctgtcggtcgcctcaag
ATRXex9_2	aaaacttgaggcgaccgacagaaacg
ATRXex9F	agtggaactgaacaagaagtgg
ATRXex9R	gaaggcacagttgataaagacacg
ATRXex9SeqF	cctgtttccctttctaattccc

Figures

Figure 4.1

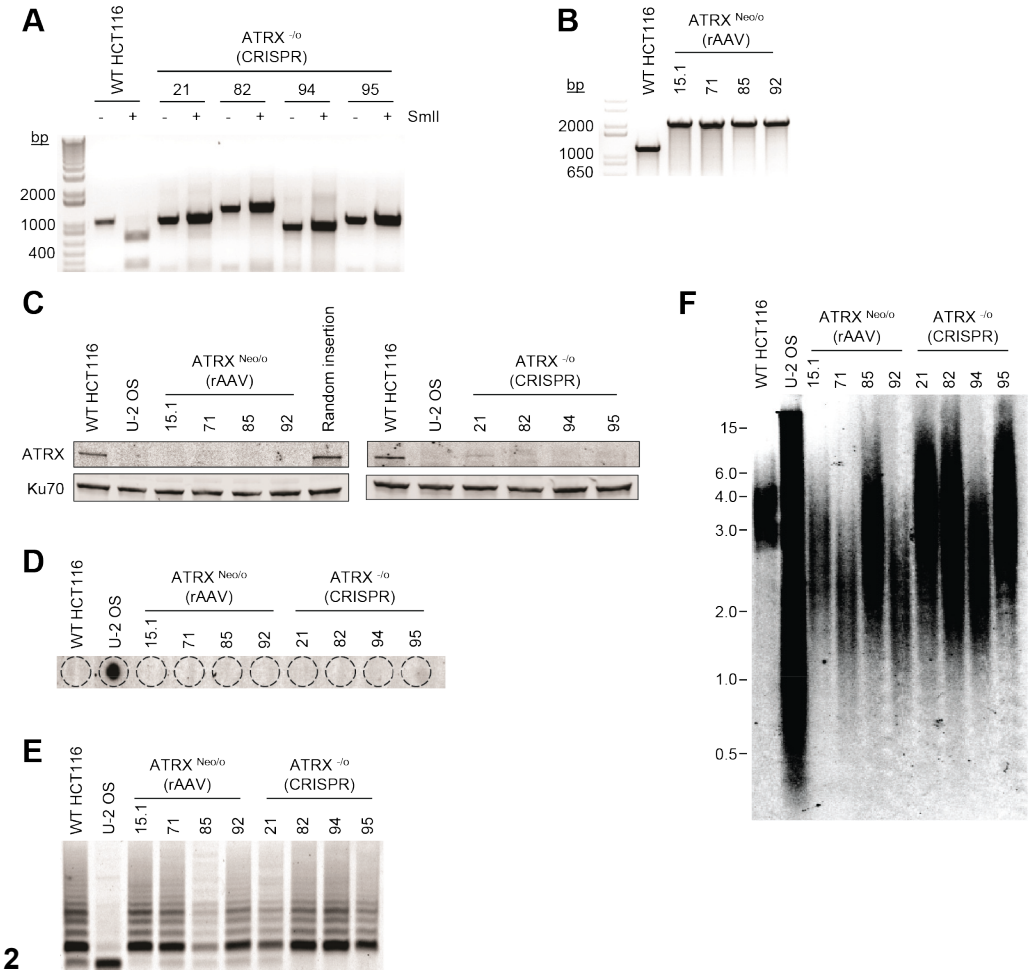


Figure 4.1. An ATRX knockout is compatible with telomerase activity.

A) PCR across the region of the *ATRX* targeted by CRISPR/Cas9 treatment in HCT116 cells shows that all clones contain modifications that result in disruption of the *Sml*I restriction enzyme recognition sequence compared to wild-type (WT). Note that *ATRX* is located on the X chromosome and HCT116 is derived from a male, making the gene hemizygous. B) PCR confirming correct gene targeting by rAAV using ATRX IntF and

ATRX IntR primers (Table 4.1). Wild-type cells should show a band of 1217 bp, while correctly targeted clones yield of band of 2186 bp. C) Western blot analysis was conducted to confirm the lack of ATRX expression in all CRISPR/Cas9 and rAAV correctly-targeted HCT116 clones. Ku70 was used as a loading control. Telomere maintenance status was analyzed (D, C-circle, E, TRAP and F, TRF assays) in both wild-type and ATRX-knockout HCT116 cells. DNA from U-2 OS cells was used as a positive control for the C-circle and TRF assays and a negative control for the TRAP assay. All HCT116 clones were telomerase-positive/ALT-negative by these criteria.

Figure 4.2

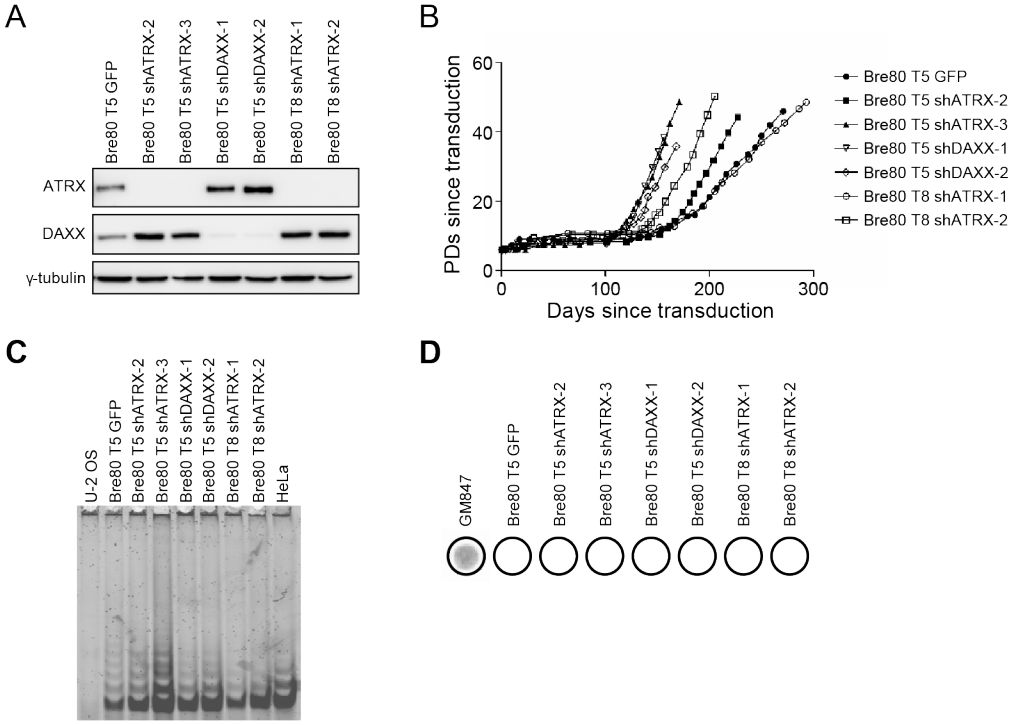


Figure 4.2. Depletion of ATRX does not induce ALT in epithelial cells.

A) The indicated Bre80 epithelial cultures were analyzed for ATRX and DAXX protein expression at a time point after immortalization; γ-tubulin was used as a loading control. B) Growth curves of each immortalized Bre80 T5 and Bre80 T8 cell line. The number of days in culture since transduction is indicated on the x-axis, while the y-axis indicates the number of PDs the culture has undergone post-transduction. C) Telomerase activity was assessed in each immortal culture using the TRAP assay. U-2 OS and HeLa cell lysates were used as negative and positive controls, respectively. D) The presence of C-circles in each immortal Bre80 T5 and Bre80 T8 cell line was determined. DNA from GM847 cells was used as a positive control.

Figure 4.3

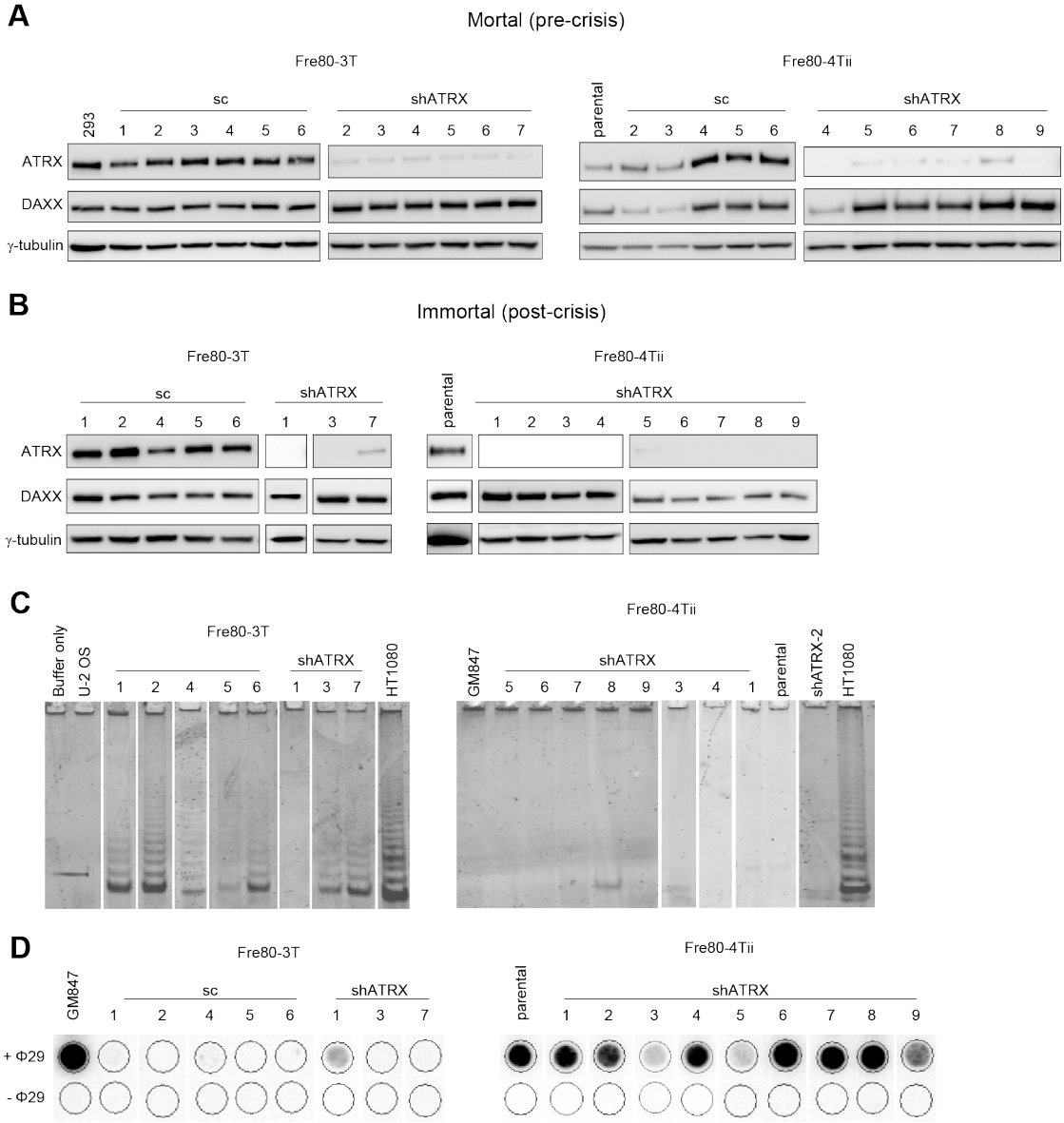


Figure 4.3. ATRX loss promotes ALT activation in breast fibroblasts.

A) ATRX and DAXX protein expression was analyzed in mortal (pre-crisis) Fre80-3T and Fre80-4Tii fibroblasts transduced with the indicated vectors, as well as untransduced parental Fre80-4Tii cells. 293 cells were included as a positive control and γ-tubulin was

used as a loading control. B) All immortal cultures were assessed for the levels of ATRX and DAXX protein expression by Western blot. γ-tubulin was used as a loading control and HT1080 cells were used as a positive control. C) All immortal fibroblasts lines were examined for the presence of telomerase activity using the TRAP assay. GM847 and U-2 OS cells were used as negative controls and HT1080 cells used as a positive control. D) The presence of C-circles was determined in each immortal cell line. GM847 cells were included as a positive control. Samples with C-circle levels above background (-Φ29 negative control) were regarded as ALT-positive.

Figure 4.4

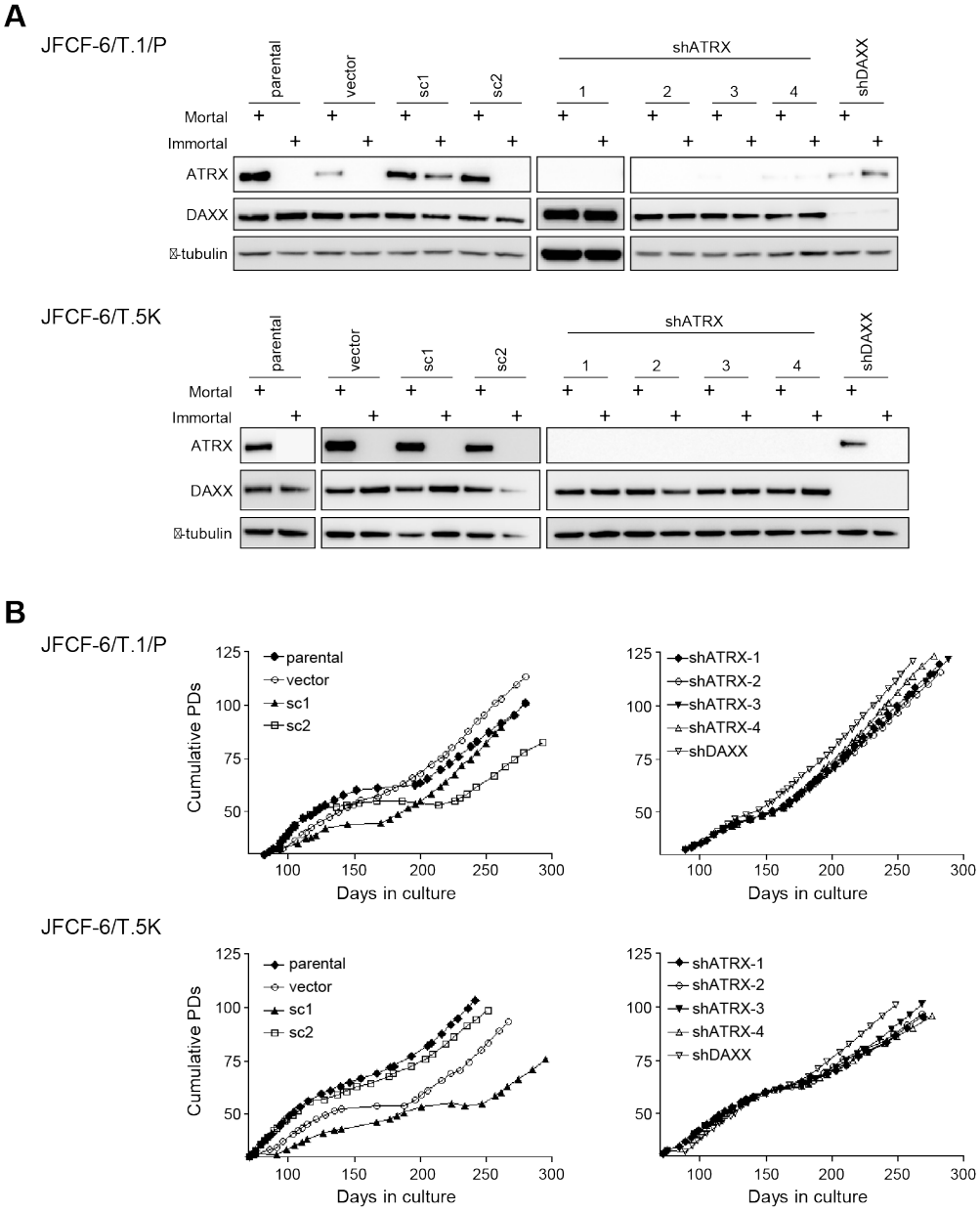


Figure 4.4. Spontaneous loss of ATRX during immortalization.

A) JFCF-6/T.1/P and JFCF-6/T.5K cells that were unmodified (parental) or transduced with an empty vector (vector), a scrambled shRNA control (sc), shATRX or shDAXX were analyzed by Western blot for the expression of ATRX and DAXX proteins. All cultures

were analyzed at a mortal and immortal time point, as indicated by the “+” above each lane. γ-tubulin was used as a loading control. B) Growth curves of each immortalized JFCF-6/T.1/P and JFCF-6/T.5K cell line. The days in culture and cumulative PDs were calculated subsequent to SV40 transformation.

Figure 4.5

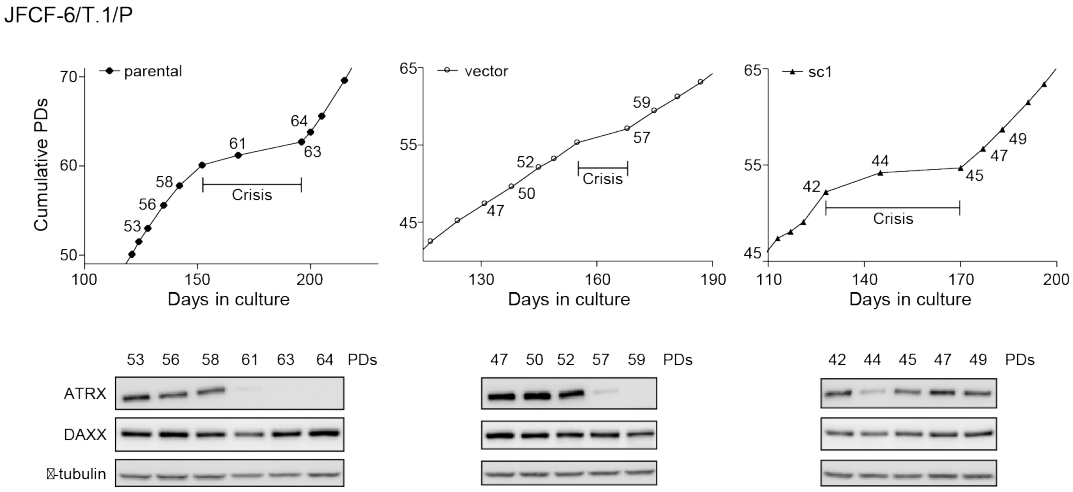


Figure 4.5. ATRX loss corresponds to a period of growth crisis.

ATRX and DAXX protein levels of three immortal JFCF-6/T.1/P cell lines (parental, vector and sc1) were analyzed at multiple PDs indicated on the corresponding growth curves; crisis was defined as the cell culture undergoing less than 1 PD/7 days. γ-tubulin was used as a loading control.

Figure 4.6

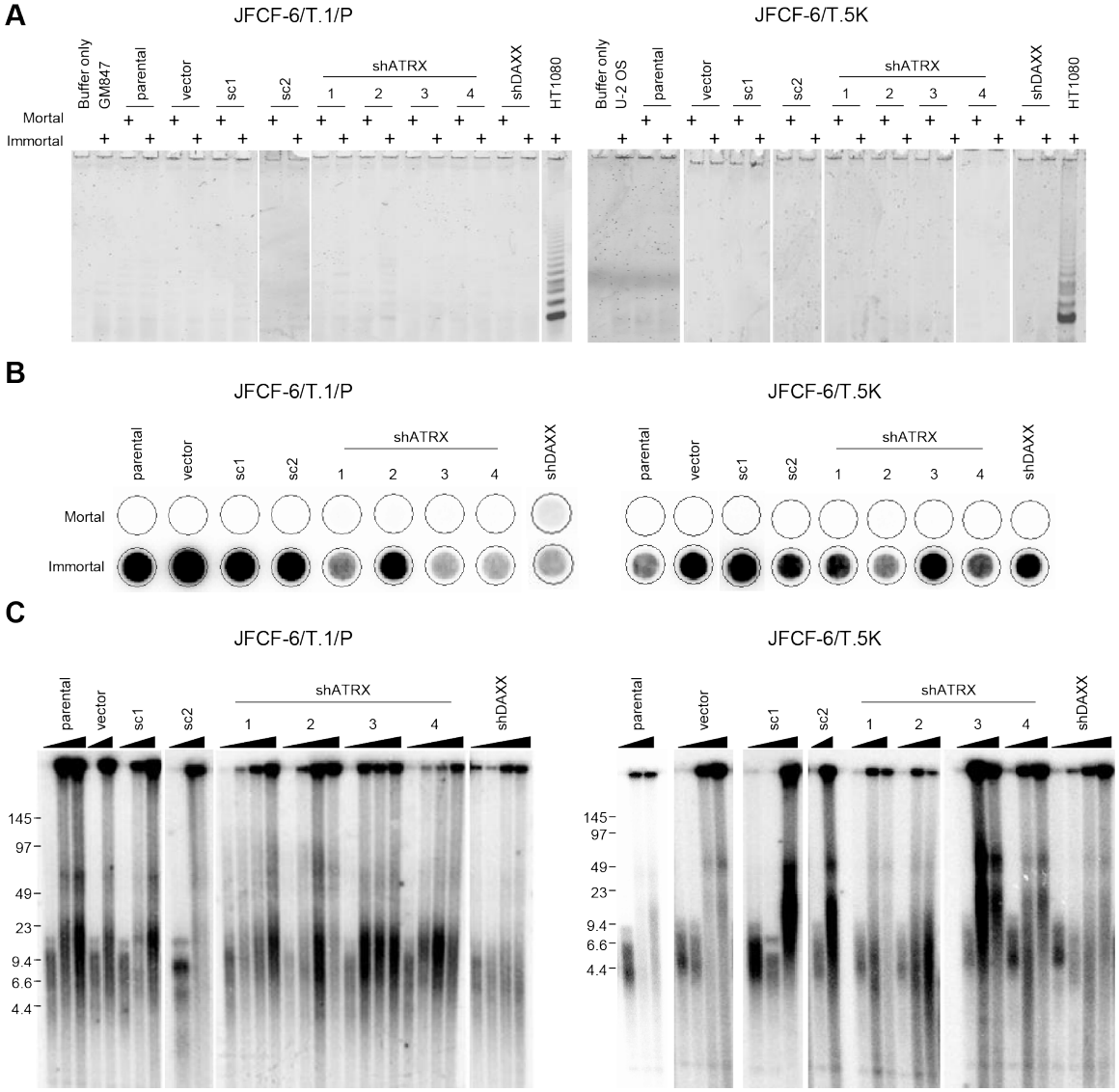


Figure 4.6. Loss of ATRX promotes fibroblast immortalization.

TEM status of JFCF-6/T.1/P and JFCF-6/T.5K cell lines treated as described for Figure 4 was analyzed. A) Telomerase activity was assessed using the TRAP assay. Each sample was assessed at both mortal and immortal time points, indicated by the "+" above each lane. GM847 or U-2 OS cell lysates were used as negative controls and HT1080 served

as a positive control. B) The presence of C-circles was assessed at both mortal and immortal time points in each cell line. C) Mean telomere length analysis using the TRF assay was performed on each cell line for at least two time points (mortal and immortal). The triangle above each set of samples indicates increasing PDs.

Figure 4.7

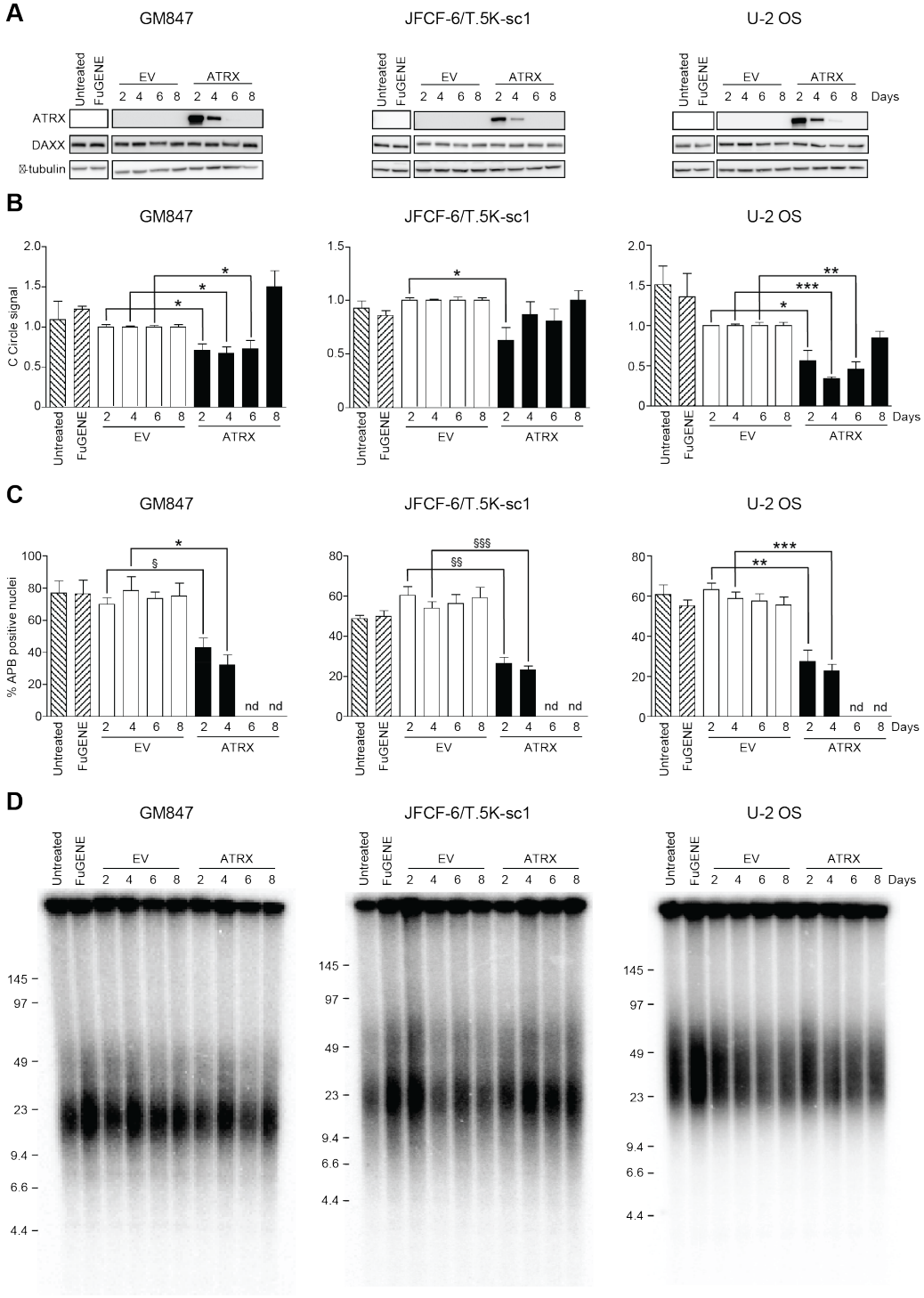

A
GM847
JFCF-6/T.5K-sc1
U-2 OS
Untreated
FuGENE
EV
ATRX
2 4 6 8 2 4 6 8
Days
ATRX
DAXX
α-tubulin
B
C Circle signal
Untreated
FuGENE
EV
ATRX
C
% APB positive nuclei
nd nd
D
145
97
49
23
9.4
6.6
4.4

Figure 4.7. ATRX expression represses the ALT mechanism.

A) ATRX and DAXX protein analysis of untreated control, FuGENE-treated, empty vector (EV)- or ATRX-transfected cultures at days 2, 4, 6 and 8 post-transfection in GM847, JFCF-6/T.5K-sc1 or U-2 OS cells. γ-tubulin was used as a loading control. The blots shown are representative of at least three separate transfections. B) The level of C-circles was assessed in each untreated control, FuGENE-treated, EV- or ATRX-transfected GM847, JFCF-6/T.5K-sc1 or U-2 OS cell culture at 2, 4, 6 and 8 days post-transfection. C-circle levels were normalized to the quantity of DNA used for each reaction, followed by normalization to the EV-transfected control at the relevant day post-transfection. Bars indicate the mean ± SEM; n = 3. * $p < 0.05$, ** $p < 0.005$ and *** $p < 0.0001$. C) APB-positive nuclei were quantified in untreated, FuGENE-treated and EV- or ATRX-transfected GM847, JFCF-6/T.5K-sc1 or U-2 OS cells at the indicated day post-transfection. A nucleus was scored APB-positive when TRF2 and PML colocalized; ATRX/APB-positive nuclei also showed nuclear ATRX staining. The bars represent the mean ± SEM; n = 3 to 5; at least 100 nuclei were counted in the untreated, FuGENE-treated or EV-transfected cultures and at least 100 ATRX-positive nuclei were counted in the ATRX-transfected cultures. D) Mean telomere length analysis was determined by TRF analysis in GM847, JFCF-6/T.5Ksc1 or U-2 OS cell cultures treated as indicated above each lane. * $p < 0.05$, § $p < 0.01$, ** $p < 0.005$, §§ $p < 0.001$, *** $p < 0.0005$, §§§ $p < 0.0001$. nd = not determined due to lack of ATRX-positive nuclei.

Acknowledgements

We thank Drs. Hilda Pickett, Tracy Bryan and Tony Cesare (CMRI) for helpful discussions and comments on the manuscript. Financial support: Cure Cancer Australia Foundation Project Grant 1062240 (to C.E.N.), Cancer Council New South Wales (NSW) Program Grant PG11-08 (to R.R.R.), National Health and Medical Research Council of Australia Project Grant GNT1009231 (to R.R.R.), National Cancer Institute Grant CA154461 (to E.A.H.), National Institutes of Health GM088351 (to E.A.H.).

Chapter 5

ATRX suppresses telomeric recombination but Its loss is not sufficient for ALT immortalization

Adam Harvey[1], Rhiannon E. Robinson[2], Julia W. Grimstead[2], Kez Cleal[2], Christine E. Napier[3], Thanh Nguyen[1], Corey Wheelock[1], Roger R. Reddel[3], Duncan M. Baird[2*], and Eric A. Hendrickson[1*#]

[1]*Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota Medical School, Minneapolis, MN 55455, USA*

[2]Division of Cancer and Genetics, School of Medicine, Cardiff University, Heath Park, Cardiff, CF14 4XN, United Kingdom

[3]Cancer Research Unit, Children's Medical Research Institute, The University of Sydney, Westmead, NSW 2145 Australia

*Co-senior authors.

[#]Corresponding author.

Author Contributions:

- Harvey, A. was responsible for the work presented in Figs. 5.1, 5.2, 5.3, 5.10 and Table 5.1. All experiments were conceptualized by Adam Harvey with assistance from Hendrickson, E. A. and collaborators.
- Robertson, R. E. and Grimstead, J. W. performed the STELA in Figs. 5.4, 5.5, 5.6, and 5.8.
- Cleal, K. performed all bioinformatics and sequencing.
- Baird, D. and Hendrickson, E. A. provided conceptual guidance.

Synopsis

The most consistent requirement of malignant progression is an escape from telomere crisis. This escape requires the activation of a telomere maintenance mechanism – either telomerase expression or the Alternative Lengthening of Telomeres (ALT) pathway – which confers cellular immortality. Whilst the mechanism of telomerase-mediated telomere maintenance is relatively well understood, the molecular requirements for the initiation of ALT are less clear. To date, one of the strongest candidate ALT-regulatory genes is ATRX, which is almost always (>90%) mutated in ALT cancers. Here, we demonstrate that the loss of ATRX enhanced by four orders of magnitude the frequency of ALT immortalization in transformed, but non-immortalized, human cells. In contrast, we show that telomerase-positive, ATRX-null cells, which are forced to undergo a second telomere crisis are only able to initiate ALT-like telomere lengthening but cannot achieve an immortal ALT-state. Interestingly, however, many of these events were length and chromosome specific, implying the presence of *cis*-acting determinants. Our data provide the first definitive distinction between ATRX's molecular control of ALT-activity and the cellular program of an immortalized ALT-cancer.

Introduction

Telomeres are repetitive DNA-protein elements that protect the ends of linear chromosomes and prevent their recognition as a double-stranded DNA break (O'Sullivan and Karlseder 2010, Napier et al. 2015). As a consequence of the end-replication problem telomeres shorten every successive cell cycle and such shortening ultimately limits the proliferative capacity of cells by eliciting a p53-dependent G_1/S cell cycle arrest that acts as a stringent tumor suppressive mechanism (Olovnikov 1973). In the absence of a functional cell-cycle arrest, continued cell division and telomere erosion ultimately results in a period of telomere dysfunction, referred to as crisis. Telomere dysfunction during crisis leads to large-scale genomic rearrangements that can drive clonal evolution and tumour progression (Lin *et al.* 2010, Roger *et al.* 2013), from which cells can escape by activating a TEM. The majority of malignancies, as well as stem cells, germ cells and single-celled organisms, almost exclusively utilize an enzyme complex called telomerase as their primary TEM (Hanahan and Weinberg 2000, Ju and Lenhard Rudolph 2008). However 10% of malignancies, predominantly those of mesenchymal origin (Henson and Reddel 2010), do not express telomerase, but instead maintain their telomeres via the ALT mechanism (Shay *et al.* 2012, Pickett and Reddel 2015).

ALT cancers enable telomere extension via a HDR-dependent mechanism (Cesare and Reddel 2010). This has been confirmed by experiments where a molecular tag was inserted into telomeres, and its duplication via HDR was monitored throughout the genome (Dunham et al. 2000). Accordingly, ALT cells have recently been found to be sensitive to inhibitors of an HR protein, ATR (Flynn *et al.* 2015). This dependence on HDR template copying between telomeres is coincident with extra-chromosomal, circular telomeric DNA

intermediates, termed C-circles, which are a strong marker of ALT activity (Henson et al. 2009), and are perhaps mechanistically involved in the HDR-dependent telomere elongation. In addition, ALT cells have been historically characterized by their lack of telomerase expression and extremely heterogeneous telomere length profiles that are likely to be derived from unregulated hyper-recombination between telomeres (Bryan et al. 1995).

More recent work has uncovered a strong correlation between ALT cancers and mutations in *ATRX* (Heaphy et al. 2011, Lovejoy et al. 2012). ATRX is a chaperone for histone H3.3, and along with DAXX, is responsible for H3.3's replication-independent incorporation into the genome (Goldberg et al. 2010, Wong et al. 2010). Specifically, H3.3 deposition into telomeric regions seems to be altered in ALT cells and, correspondingly, telomeric chromatin dynamics are altered (Goldberg et al. 2010, Episkopou et al. 2014, Ivanauskiene *et al.* 2014, O'Sullivan and Almouzni 2014). Because *ATRX* is mutated so frequently in ALT tumors this implies that ATRX's chromatin regulation of telomeric chromatin may normally be repressive of the HR-activity observed in ALT cells. How ATRX might contribute to the onset of ALT is still unknown (Amorim *et al.* 2016). The mutation or deletion of *ATRX*, however, does co-segregate with the ALT phenotype in cell fusion experiments (Bower et al. 2012). Additionally, the frequency of ALT immortalization events increases with shRNA knock-down of ATRX in human fibroblasts (Napier et al. 2015, Hu *et al.* 2016). Lastly, ectopic re-expression of ATRX in ALT cells diminishes the ALT activity of the cells (Clynes et al. 2015, Napier et al. 2015).

Here, we report for the first time, direct genetic evidence that the loss of *ATRX* initiates ALT activity. We demonstrate using genetic knockout experiments of *ATRX* in transformed, non-immortalized cells that correct gene targeting correlates 100% with

subsequent ALT immortalization. Furthermore, we have devised a model system that, for the first time, is able to demonstrate the earliest effects of *ATRX* loss on cells undergoing telomere crisis.

Results

Loss of ATRX in transformed, non-immortal cells enhances the frequency of immortalization

We previously generated *ATRX*-null telomerase-positive human HCT116 cells by both rAAV- and CRISPR/Cas9- mediated gene targeting and demonstrated that the genetic deletion of *ATRX* did not lead to the activation of the ALT phenotype (Napier et al. 2015). Thus, the loss of ATRX activity in a tumor cell that is already utilizing telomerase as its primary telomere maintenance mechanism is, *per se,* not sufficient to trigger a switch to ALT. We additionally showed that knockdown of ATRX by siRNA increased the frequency of ALT immortalization, but still required a natural selection process through telomere crisis where presumably the vast majority of cells with diminished ATRX levels did not initiate ALT. This result implied that either the initiation of ALT was very sensitive to the levels of ATRX protein on a cell-by-cell basis, or that a second genetic or epigenetic event occurs during telomere crisis that is required for ALT immortalization. The latter of these possibilities seemed more likely, as the genetic deletion of *ATRX* in a telomerase-positive cell did not initiate ALT. To address this question, we initiated a genetic knockout study of *ATRX* in a more pathophysiologically relevant human cell culture system by using a SV40-transformed primary human cell line, JFCF-6/T.1P (henceforth JFCF). Importantly, while these cells are oncogenically transformed, they are not immortalized, and thus closely resemble human cancer cells as they pass the Hayflick limit and move into telomere crisis (Yeager et al. 1999). We reasoned that *ATRX*-mediated gene targeting in these cells would enable us to elucidate the genetic contribution of *ATRX* in repressing ALT. JFCF cells can spontaneously immortalize in culture, but do so at a low frequency

of 10^{-7}, and notably, can do so by utilizing either the telomerase or ALT pathways (Sprung *et al.* 1997, Yeager et al. 1999). For our knockout approach, we employed rAAV-based gene targeting where exon 5 of the *ATRX* gene was replaced with a *NEO* drug resistance marker, according to previous rAAV-mediated targeting strategies (Leung et al. 2013, Napier et al. 2015) (Fig. 5.1A). After three independent infections, drug resistant, single-cell clones were propagated for about 2 months. Across all three infections after two weeks there were hundreds of clones alive (n = 388), however after 2 months only 16 had survived and these were growing as if they were immortal (Fig. 5.1B, Table 5.1). We confirmed their immortality by continuing to culture the clones for an additional 3 weeks with no indication of a growth defect (Fig. 5.1C), and were able to grow them in culture beyond 50 population doublings (PDs) for 3 months post-immortalization. As 16 of 388 original single-cell clones had immortalized, we conclude that the rate of immortalization in these experiments is ~4 X 10^{-2}, which is between 10^4- to 10^5-fold higher than the spontaneous rate of immortalization for JFCF (Sprung et al. 1997, Yeager et al. 1999). In order to assess whether these immortal clones coincided with correct gene targeting, we utilized PCR and confirmed that 100% of the clones were correctly targeted (Fig. 5.1D). This high frequency of correct targeting was striking as knockouts in previous studies with HCT116 had yielded targeting efficiencies of only ~4% (Napier et al. 2015). Presumably the difference between targeting efficiencies in these cell types was due to the selection for immortalization in the case of the JFCF cells undergoing crisis. To confirm this presumption, these PCR results were substantiated by Western blot analyses, which demonstrated that none of these clones expressed ATRX protein (Fig. 5.1E) in contrast to the parental JFCF cells, which did (Fig. 5.2). Importantly, we also carried out a parallel experiment in which an irrelevant gene - phosphatidylinositol glycan anchor biosynthesis

class A (*PIGA*) - was targeted with a nearly identical vector. Similarly, 173 drug-resistant clones were isolated after two weeks, but in this instance none of the clones immortalized (Table 1). From these experiments, we concluded that the loss of *ATRX* correlated with immortalization in JFCF cells in 100% (16/16) of the clones analyzed and occurred at a frequency 4^{+} orders of magnitude higher than can occur spontaneously. Importantly, the occurrence of crisis (Fig. 5.1C) demonstrates definitively that at least one additional step, in addition to complete ATRX loss, is required for immortalization.

Loss of ATRX before telomere crisis is required for ALT immortalization

The 16 immortalized clones were subsequently categorized for their telomere elongation maintenance mechanism. TRAP analysis revealed that all of the immortal clones were telomerase negative (Fig. 5.3A), which by itself strongly suggested that these cells were ALT. Second, by terminal restriction fragment (TRF) analyses all the clones had elongated ALT-like telomere lengths (Fig. 5.3B) (Bryan et al. 1995). Third, all of the clones were positive for the presence of C-circles (although some were more strongly positive than others), a hallmark of ALT activity (Fig. 5.3C) (Henson et al. 2009). A fourth characteristic of ALT cells is the incidence of the APBs (Yeager et al. 1999). These nuclear protein conglomerates are postulated to be the molecular scaffolding required for the HR-dependent telomere lengthening characterized by the ALT-phenotype (Conomos et al. 2014). Accordingly, we found a striking co-localization phenotype between large telomeric foci with the PML protein only in the *ATRX*-null clones (Fig. 5.3D-F). *In toto*, these data strongly suggest that in cells undergoing telomere crisis the absence of ATRX results in the onset of ALT.

The absence of ATRX in telomerase-positive cells can induce ALT activity when combined with telomere stress

After establishing a direct link between the absence of ATRX in cells undergoing telomere crisis and ALT activation, we sought to understand if this event was limited to cells which had yet to commit to a TEM pathway. Given that we had shown that knockout of *ATRX* in telomerase-positive cells did not yield any evidence of ALT activity, we now asked whether inducing telomere-crisis in these cells would result in activation of ALT. To answer this question, we utilized our *ATRX*$^{Neo/o}$ HCT116 cells (Napier et al. 2015) and expressed DN-hTERT in them to abrogate telomerase activity and induce telomere erosion. We have validated this system in WT HCT116 cells (Hahn et al. 1999, Jones et al. 2014) where the expression of DN-hTERT led to telomere erosion, fusion and crisis. All (100%) WT HCT116 cells, however are ultimately able to escape crisis by achieving the reactivation of telomerase activity (Jones et al. 2014). In the case of *ATRX*$^{Neo/o}$ HCT116 cells the expression of the DN-hTERT also led to an abrogation of telomerase activity (Fig. 5.4A). A total of 33 clones were picked and their growth was monitored; in contrast to WT HCT116 clones only 5 (15%) of the *ATRX*$^{Neo/o}$ DN-hTERT clones were able to escape crisis, with 28 (85%) of these clones failing to escape crisis despite monitoring for up to 220 days at which point no cells were left in the cultures (Fig. 5.4B). Thus, surprisingly, these data indicated that, in contrast to JFCF cells (where the absence of ATRX greatly increased the frequency of immortalization), the absence of ATRX compromises the ability of HCT116 cells to escape a telomere-driven crisis.

Of the five clones that escaped crisis, 3 escaped rapidly (demarcated in black, Fig. 5.4B), resuming a normal rate of growth after approximately 60 days from the point of single cell cloning, whereas the two remaining clones escaped crisis with slower kinetics

(demarcated in blue, Fig. 5.4B), resuming a normal rate of growth after 110 days. We confirmed that these clones were genetically distinct by next generation sequencing, where unique SNPs in each clone were validated (Fig.5.5A). An unbiased phylogenetic algorithm was used to determine that the clones were not all from an identical lineage (Fig.5.5B). We monitored the telomere length of these clones over time using single telomere length analysis (STELA) of the XpYp and 17p telomeres (Baird et al. 2003). All of the clones analyzed displayed telomere erosion in the first sampling points prior to crisis, consistent with gradual end-replication losses following the abrogation of telomerase activity. The clones that were not capable of escaping crisis displayed gradual telomere erosion all the way to the final sampling point (clones #1, #5, #8, #11, and #37; Fig. 5.4C and Fig.5.6). Similarly, the two clones that escaped crisis with slower kinetics, displayed gradual telomere erosion that ultimately stabilized at a short, but apparently functional, length (clones #31 and #32; Fig.5.7). The stable telomere length in these clones was consistent with the fact that both clones ultimately re-expressed significantly amounts of telomerase activity (demarcated in blue, Fig. 5.4A).

In complete contrast, the three clones that underwent a rapid escape from crisis, displayed large and precise telomeric elongation events at both the 17p and XpYp telomeres (clones 2, 3 and 4, Fig. 5.4C). The elongation events appeared to occur only on the shorter of the two telomeric alleles observed at 17p. Intriguingly, the telomeric elongation events were apparently identical in all three clones and were specific to each telomere with extension to a mean of 6.8 kb at 17p and an extension to a mean of 2.1 kb at XpYp. The rapidity (within several PDs) and extent (up to an additional 5 kb) of telomere extension was not consistent with telomerase addition, but was consistent with HR-mediated processes.

We considered that the bimodal telomere length distributions observed with STELA represented either sub-clonal populations derived from a single cell that had exhibited the telomeric extension, or that the extension event had occurred at a single telomeric allele and that the distributions were consistent with a population displaying one elongated allele and one un-altered allele. To distinguish between these possibilities, we undertook single cell subcloning of clone 3 at PD31 (Fig. 5.4C) and analyzed the telomere length profiles in the subclones after 17 to 22 PDs. Of the nine subclones analyzed, all apart from one displayed bi-modal telomere length distributions at 17p consistent with the presence of one elongated and one un-altered telomeric allele (Fig. 5.8A). The single subclone that was only able to survive 17 PDs before dying (clone 2, Fig. 5.8A), displayed a unimodal telomere length distribution and no evidence of telomeric elongation.

We have observed that the escape from a telomere-driven crisis, in the presence of functional ATRX, is accompanied by large-scale genomic rearrangements that can result in the amplification of the hTERT locus.(Jones et al. 2014) We considered that if the initiation of an ALT-like phenotype arose because of the induction of aberrant HR activity at telomeres, this may also lead to more widespread mutation across the genome. We therefore undertook whole genome sequencing at the same PD points at which telomeric elongation was observed in the three clones. All three genomes displayed relative stability, with only one clone exhibiting a single deletion at the 5p chromosome arm (Fig. 5.9). Thus, whilst the escape from crisis in the absence of ATRX can be associated with rapid, telomere-specific elongation, this does not appear to be accompanied by large-scale genomic copy number changes.

Taken together these data indicate that the absence of ATRX compromises the ability of HCT116 cells to escape a telomere crisis. Of those clones that did escape, two of the

escapees (clones #31 and #32) appeared to utilize telomerase whereas three escapees (clones #2, #3 and #4) appeared to (at least initially) engage an ALT-like elongation process. This elongation process was highlighted by rapid and precise chromosome-specific telomeric extension of the shorter telomeric alleles and this did not have a significant impact on genome stability.

ATRX represses ALT activity, but its loss is not sufficient for sustained ALT activation

We now considered whether the maintenance of ALT was also ATRX-dependent by undertaking long-term culture to >70 PDs (>200 days) and analyzing telomere dynamics and telomerase activity in the five clones that escaped crisis (Fig. 5.10a, Fig.5.7 and Fig. 5.11), as well as one sub-clonal population (clone #3, sub-clone 11; Fig. 5.8). The two clones that displayed a slower escape from crisis (demarcated in blue, Fig. 5.4) showed an increase in telomerase activity (demarcated in blue, Fig. 5.4A) and a stabilization of telomere length (Fig. 5.7) consistent with that observed in clones transiting a telomere-driven crisis following the up-regulation of telomerase activity (Jones et al. 2014). A clear pattern of telomere dynamics emerged in the clones in which the telomere elongation event had occurred (clones demarcated in black, Fig. 5.4B); in these clones the extended telomeric allele exhibited gradual telomere erosion, with ongoing cell division, at rates that were indistinguishable from that observed in the absence of telomerase (mean 90 bp/PD (Baird et al. 2003); Fig. 5.8B, Fig.5.10A, Fig.11). In contrast, the rate of telomere erosion of the shorter un-mutated alleles was less (mean 23 bp/PD) with lengthening of this allele in later passages such that the two alleles homogenized and appeared to stabilize (Fig. 5.8B, Fig. 5.10A, Fig.5.11). The specific elongation events at the 17p and XpYp telomeres

were reflected in the genome-wide telomere length distributions, where an increase in general telomere length and heterogeneity was observed in these clones via TRF; however both the telomere length and heterogeneity decreased with ongoing cell division in stark contrast to that observed in long-term ALT-positive cell cultures (Fig. 5.10B). Consistent with the observed telomere dynamics, the telomerase activity increased with ongoing cell division (clones demarcated in black, Fig. 5.4A) being virtually undetectable at PD30, but rising to 25% to 77% of the *ATRX*$^{Neo/o}$ HCT116 parental cell's level after PD80^{+}. These data are consistent with a model where the absence of ATRX permits critically short telomeres to engage in an ALT-like recombination pathway but that this state cannot be stably maintained. Over time, the telomeres begin to erode and ultimately the survival of the clones is dependent upon the reactivation of telomerase. Consistent with the view that telomerase and ALT may be competing for telomere maintenance in these cells — with telomerase ultimately gaining the upper hand — we observed variable amounts of C-circles in these clones throughout the experiment but with the lowest levels observed at the latest time points (Fig. 5.10C, D).

These data demonstrate that ATRX suppresses telomere elongation events, which are reminiscent of ALT, but that the long-term maintenance of this state in HCT116 cells is suppressed by some other factor. Instead, following the initial telomeric lengthening in HCT116 cells, progressive telomere erosion of the elongated alleles ensues and continued survival occurs when telomerase is subsequently upregulated. This leads to the preferential lengthening of the shorter telomeric allele, the homogenization of allelic telomere length and the maintenance of replicative lifespan. This is in contrast to *ATRX*-null JFCF fibroblasts that readily maintain the ALT state.

Discussion

We and others, have established a strong link between *ATRX* and the ALT phenotype in various malignancies and cell culture models (Heaphy et al. 2011, Bower et al. 2012, Lovejoy et al. 2012, Clynes et al. 2015, Napier et al. 2015); whilst there has been a focus on the association of *ATRX* with such phenotypes or mechanistic contributions to repressing ALT activity, thus far it has not been possible to demonstrate that the loss *ATRX* alone is sufficient to induce ALT in human cells. The lack of ALT-activity in *ATRX*-null cells was surprising particularly given that the association of the lack of ATRX expression with ALT cancers is so high. Adding to this complexity was the finding that genetic ablation of another H3.3 chaperone, *ASF1*, readily generated ALT-activity in telomerase-positive cells (O'Sullivan et al. 2014). Here we present strong evidence that the loss of *ATRX* during telomere crisis activates ALT in transformed, but not-yet immortalized human cells, greatly increasing the frequency of ALT-immortalization events by >1 x 10^4. While other recent experiments have been able to genetically derive ALT cells from a common non-immortalized progenitor by functional inactivation of *TERC* (Min *et al.* 2017), the work shown here lends insight to the consequences of the *ATRX*-status as cells transit crisis. Moreover, the occurrence of subsequent periods of crisis in our clones provides definitive evidence for the first time that complete ablation of ATRX activity is insufficient to activate ALT and that an additional event(s) is required.

The efficiency with which this putative second event in JFCF cells occurs may be facilitated by the absence of functional p53 (due to the expression of the SV40 tumor antigen). In contrast, this second event did not appear to occur in *p53*-competent, *ATRX*-null HCT116 cells that were forced to undergo telomere crisis, because while ALT activity

was initiated it was not maintained. This model provides unprecedented insights into the events that are driven solely by *ATRX* loss. For the cells that were able to escape crisis by apparently initiating ALT (n = 3), we used high-resolution telomere length analysis to track the dynamics of single telomeres during the transition through crisis and were able to observe elongation events that specifically occurred on the shortest telomeres. At 17p we observed two different length telomeric alleles and following the expression of DN-hTERT both alleles were subjected to telomere erosion. We then observed a specific elongation event at 17p, that appeared to involve the shorter of two telomeric alleles. Intriguingly, all three clones (which were independent) exhibited an apparently identical telomere length change at 17p (mean of 6.8 kb). This was mirrored at the XpYp telomere, where the single detectable telomeric allele was subject to similar elongation (mean of 2.1 kb) in all three clones. The clones that escaped with telomeric elongation, all did so with very rapid kinetics (all within 10 PDs). This observation is simply not consistent with the escape of a single cell from crisis (which would require a minimum of 20 PDs to re-populate the culture dish), but would be consistent with a mechanism where multiple cells from each single clone were capable of escaping from crisis. These observations indicate that the specific elongation events occurred independently in multiple cells, within each of three independent clones. The nature of these initial elongation events and the underlying genetic determinants remain to be elucidated, but the presence of distinct elongation events specific to each telomere, strongly suggest that this process is tightly controlled. We speculate a potential role for *cis*-acting sequence elements that provide a template for elongation and which result in telomere-specific lengthening events. It is pertinent to note that “Template for ALT” regions that integrate into telomeres have been characterized in telomerase-null *C. elegans* survivors (Seo *et al.* 2015). Moreover, orphan nuclear receptor

NR2C/F-mediated bridging may be responsible for bringing telomeric and non-telomeric DNA into proximity to mediate *cis*-acting rearrangements (Marzec *et al.* 2015). Lastly, it should be noted that in our experiments we detected almost exclusively telomere-specific elongation events accompanying the loss of *ATRX*. In contrast, in a mouse model for glioblastoma a reduction of ATRX was correlated with an accompanying significant genetic instability (Koschmann *et al.* 2016). It is likely that these differences may reflect the known biological telomeric differences between human and mouse or the fact that additional genetic alterations were required to produce the glioblastomas.

Our data in DN-hTERT-expressing HCT116 cells distinguished ALT-immortalization from ALT-like telomere recombinations. We suggest that ATRX functions at the molecular level to repress ALT-mediated recombinations, as *ATRX*-null HCT116 cells exhibited ALT-activity at the time of telomere crisis escape. This is consistent with the large body of biochemical evidence that ATRX-modulates telomeric homeostasis (Goldberg et al. 2010, Wong et al. 2010, Clynes et al. 2015, Eid et al. 2015). Our results importantly imply that ALT-activity at telomeres is distinct from an ALT-state of continuous telomere recombinations. One obvious explanation for this is the considerable amount of data supporting the model that telomere proteins coordinated by the Shelterin complex, actively repress DNA damage/recombination signaling (Denchi and de Lange 2007, Palm and de Lange 2008, Poulet *et al.* 2009, Sfeir and de Lange 2012, Rai et al. 2016). Intriguingly, many of these DNA damage response pathways signal through *p53* (Karlseder *et al.* 1999, Stagno D'Alcontres *et al.* 2007, Fujita *et al.* 2010), and HCT116 is notably a WT p53-expressing cancer cell line (el-Deiry *et al.* 1994). Recent work has additionally described an ALT-specific genome histone mark, H3.3 phosphorylated at S31 (Chang *et al.* 2015), which is of significant interest as both ATRX and ASF1 are H3.3 chaperones shown to

regulate ALT. Independently, this has recently been shown to be a global histone modification that results from chromosomal mis-segregation at mitosis, and which triggers a p53-dependent cell cycle arrest (Hinchcliffe *et al.* 2016). Because ALT cells are clearly not cell cycle arrested, the implication of these two findings is that that p53 signaling must be somehow disrupted in ALT tumors, which is indeed a common observation (Henson and Reddel 2010). Another explanation for the lack of persistence of an ALT state in our HCT116 *ATRX*$^{Neo/o}$ DN-hTERT cells could be due to the presence of both DN-hTERT bound to the telomeric ends. Recent work has suggested that hTERT expression in ALT cells can alter the G-overhang on the leading and lagging strands, which was proposed to influence the ALT-recombination activity (Min et al. 2017). Thus, a DN-hTERT might alter the state of these overhangs and their recombinogenic potential and thus influence the persistence of ALT. Future work should attempt to further clarify the genetic contributions to both ALT-recombinations *per se,* and an active ALT state.

Our data indicate that the timing of telomeric elongation is dependent on telomere length and occurs when a single telomeric allele reaches a specific length threshold at which point the large-scale elongation event occurs. The definition of this threshold requires detailed clarification, however it is pertinent to note that *ATRX*$^{Neo/o}$ HCT116 cells expressing DN-hTERT contain populations of cells with telomeres of less than 400 bp, which is the point at which telomeric elongation is observed [see for example clones 2 (PD30), 3 (PD31) and 4 (PD29), respectively (Fig. 5.10 and Fig. 5.11). In contrast WT HCT116 cells expressing DN-hTERT all entered crisis with telomere length distributions that rarely extended below 400 bp because these telomeres are subjected to fusion (Jones et al. 2014). Our data imply that in the absence of ATRX, shorter telomeres are more likely to be processed for ALT-like elongation rather than telomere fusion. Moreover, our

data imply that even if telomere fusion occurs to drive genome instability in *ATRX*$^{Neo/o}$ HCT116 cells expressing DN-hTERT, the rapid elongation of the shortest telomeres provides a temporary relaxation of the stringent selective conditions for the upregulation of a telomere maintenance mechanism thereby accounting for the relative genome stability observed in cells immediately after telomeric elongation.

Finally, in this study we utilized two cellular models of malignant immortalization: SV40-transformed mesenchymal JFCF cells and epithelial colon cancer HCT116 cells. We chose these two models to synergistically complement the strengths and weaknesses of the other: JFCF cells are a more “normal” cell line, but are limited for mechanistic studies; in contrast, HCT116 cells are already immortalized and transformed, but can be utilized for very detailed mechanistic studies. Thus, while the functional inactivation of *ATRX* In JFCF cells resulted in a dramatic increase in the frequency of ALT immortalization, we could not molecularly determine the mechanism of that escape from crisis, as correctly-targeted clones could only be identified after cells had immortalized. Therefore, we can only presume that as these cells entered crisis, the absence of ATRX permitted telomeric recombinations similar to what was observed in our HCT116 studies. Interestingly, in JFCF these presumptive recombinations permitted the development of ALT whereas in HCT116 cells they did not. These differences established in our two different cell-based studies are reflected in the telomere maintenance mechanisms adopted by different tumor types, as ALT predominates in tumors of mesenchymal original (similar to JFCF) and telomerase dependence in the majority of epithelial-derived (similar to HCT116) carcinoma (Henson and Reddel 2010).

Materials and methods

Cells culture

HCT116 cells were purchased from the ATCC and maintained in McCoy's 5A media supplemented with 10% FBS, 1% glutamine, and 1% penicillin/streptomycin. JFCF-6/T.1/P cells(Napier et al. 2015) were maintained in DMEM supplemented with 10% FBS, 1% glutamine, and 1% penicillin/streptomycin. The JFCF and HCT116 cultures were both derived from males. Notably, *ATRX* is located on the X-chromosome, and thus gene targeting strategies must only target one allele. Cells were maintained in 10 cm plates and passaged every 3 to 5 days. To initiate telomere erosion in $ATRX^{Neo/o}$ HCT116 cells expressing DN-hTERT, the cells were transduced with amphotropic retroviral vectors containing a DN-hTERT cDNA (Hahn et al. 1999), as described (Preto et al. 2004).

Gene targeting and *ATRX* knockouts

The *ATRX* knockout by exon replacement with rAAV was performed as described (Napier et al. 2015). Briefly, homology arms were constructed by PCR, flanked by an IRES-Neo cassette, and ligated into an rAAV production vector. Producer 293-AAV cells were co-transfected with pAAV Helper and pAAV Rep/Cap, as described (Khan et al. 2011). JFCF cells (5×10^5) were plated approximately 24 hr prior to rAAV-infection. Cells were infected with virus-containing media, and 48 hr-post infection, the media was changed to include 0.5 mg/mL G418. Drug-resistant colonies were counted 2 weeks after infection, and allowed to continue to grow for another 2 to 4 weeks until control-infected cells had succumbed to telomere crisis. Independent clones were isolated by cloning cylinders and expanded for further analysis.

Telomere-relevant Southern blotting (TRF and C-circle analyses)

Genomic DNA was extracted from ~1 x 10^7 cells, and 50 μg of genomic DNA was digested with HinfI and RsaI, as described (Henson et al. 2009). For the TRF analyses, 12 μg of digested genomic DNA was resolved overnight on a 0.7% agarose 1 x TBE gel. This gel was depurinated, denatured, and neutralized, followed by overnight capillary transfer to a nitrocellulose membrane. For the C-circle assay, 30 ng of digested genomic DNA was incubated for 8 hr with phi29 DNA polymerase, in the buffer supplied by the manufacturer (New England Biolabs), supplemented with 1 mM dATP, dGTP, dTTP, and 0.2 mg/mL BSA with 0.1% Tween-20. Reactions were chemically denatured, neutralized, and dot-blotted onto a nitrocellulose membrane. The membranes were pre-hybridized for 1 hr with Church's buffer, then hybridized overnight in 4X SSC at 55°C. Membranes were washed 3 times with 4X SSC and once with 4X SSC + 0.1% SDS, each for 30 min, exposed to a phosphorimaging screen, and detected and quantitated with a Typhoon phosphoimager.

Immunofluorescence and telomere FISH (IF-FISH)

This assay was performed as described (Conomos et al. 2014, Napier et al. 2015). Cells (1 x 10^5) were plated on chamber slides, and allowed to grow for 24 hr. Cells were washed once with PBS, then fixed with 4% formaldehyde in 1X PBS. Blocking and RNAseA treatment (0.1 mg/mL) were performed in antibody dilution media (ABDIL; 20 mM Tris-HCl pH 7.4, 0.2% fish gelatin, 2% BSA, 0.1% Triton X-100, 150 mM NaCl, and 0.1% sodium azide) at room temperature for 30 min. Cells were stained with a PML antibody (Santa Cruz BioTech sc-9862) diluted in ABDIL for 1 hr, washed 3 times with 1X

PBS + 0.1%Tween-20 (PBST), and incubated with Alexa-488 α Goat IGG secondary antibody diluted in ABDIL for 1 hr. Cells were washed in PBST, fixed with 4% formaldehyde and prepared for FISH hybridization. Telo-C (PNA Bio) PNA probe was hybridized to the slides at 80^{o} in hybridization buffer (10 mM Tris-HCl pH 7.4, 4 mM Na_2HPO_4, 0.5 mM citric acid, 1.25 mM $MgCl_2$, 0.25% blocking reagent (Roche) and 70% formamide). Slides were washed, counterstained with DAPI, and mounted with ProLong Gold (ThermoFisher). Microscopy was performed with a Nikon-TiE deconvolution bright-field microscope with a 60X objective.

Telomerase activity

Telomerase activity was detected using the TRAPeze XL Telomerase detection kit (Chemicon international). Cell pellets were lysed in 200 μL CHAPS buffer. Each lysate was included in the TRAPeze reaction using the Amplifluor primers and Titanium Taq (BD Clontech). Each experiment included positive control extracts, control template for the generation of a standard curve, heat-treated controls, no template control and an internal control to monitor PCR inhibition. Telomerase activity was expressed as TPG (Total Product Generated) values. Each unit of TPG corresponds to the number of TS primers extended at least 3 telomeric repeats by telomerase.

Telomere Length Analysis (STELA)

DNA was extracted from using standard proteinase K, RNase A, phenol/chloroform protocols. For telomere length analysis we used the single telomere length analysis (STELA) assay as described (Baird et al. 2003, Capper *et al.* 2007). The XpYp telomere

of HCT116 has one telomeric allele that is detectable using the XpYpC primer (5'-CAGGGACCGGGACAAATAGAC-3 ') for the STELA reactions.

Whole genome sequencing and copy number analysis

Paired-end sequencing was performed on the Illumina HiSeq x10 platform to a median depth of 15X and a read length of 150 bp. Reads were mapped to hg19 (2009) using bwa-mem (version 0.7.12) (Li and Durbin 2009). To estimate relative copy number profiles, sequencing coverage was calculated over 10 kb non-overlapping windows and normalized to the median genome coverage of each sample. De-noising of coverage values was then performed using a wavelet regression method through use of the PyWavelets package (available online at: http://www.pybytes.com/pywavelets). Here, coefficients were obtained by decomposing the data using a multilevel discrete wavelet transform with a Haar wavelet (functional parameters were level = 2, mode = 'per') (Lio 2003). Denoised coverage values of each chromosome were then obtained by applying an inverse discrete wavelet transform over the thresholded coefficients, with the threshold defined by: $\lambda=\sigma\sqrt{2\log(N)}$ where σ is the median absolute deviation and N is the number of data points. To determine the relative copy number change profile of each sample with respect to other samples, the "parental" copy number profile common to all samples was inferred by first excluding outlier bins. At each genomic bin, outliers were detected as points with a median absolute deviation > 2. The "parental" copy number was then calculated as the median of inlier bins, and the relative copy number change profile for each sample was determined as the denoised coverage values - parental copy number profile.

Unique SNP identification

To identify the unique SNPs in our cohort, candidate variants were first identified using the Genome Analysis Tool Kit (GATK v3.3) according to best practice guidelines (McKenna *et al.* 2010, DePristo *et al.* 2011). Variants with a genotype quality >30 were kept whereas multiallelic and indel sites were discarded. Variant locations were then intersected using BedTools to remove common call sites (Quinlan and Hall 2010). Additionally variants which overlapped simple repeat intervals as documented in the UCSC simple repeat track were discarded (Benson 1999, Rosenbloom *et al.* 2015). Due to the potential limited sensitivity of GATK for SNP calling, a further filtering step was conducted to remove sites with common variant allele mappings across more than one sample. Here the alignment files for each sample were crosschecked using Pysam for each candidate variant . If any reads could be identified in more than one of the samples which supported the variant allele, that allele was discarded.

Phylogenetic reconstruction

Using the table of unique SNPs generated for each sample, a distance matrix $M = (d_{ij})$ was calculated according to $d_{ij} = (x_i + x_j) / 2$ where x_i, x_j are the number of unique SNPs for samples *i* and *j* respectively. Tree reconstruction was performed using the neighbor joining algorithm implemented using the PyCogent software package (Knight *et al.* 2007).

Acknowledgments

Funding: This work was funded, in part, by a grant from the National Institutes of Health (GM088351), grants from the National Cancer Institute (CA154461 and CA190492), a seed grant provided by the Karen Wyckoff Rein in Sarcoma Foundation administered through the University of Minnesota Cancer Center, a grant from Cancer Research UK (C17199/A18246), and a grant from the National Health and Medical Research Council of Australia (GNT1088646). **Author contributions:** A.H., C.E.N., R.R.R, D.M.B. and E.A.H. designed the study and wrote the manuscript. A.H., C.E.N., R.E.R, J.W.G., K.C., T.N., and C.W. performed experiments. **Competing interests:** E.A.H. declares that he is a member of the scientific advisory boards of Horizon Discovery, Ltd. and Intellia Therapeutics, companies that specialize in applying gene editing technology to basic research and therapeutics. **Data and materials availability**: All data needed to evaluate the conclusions in this paper are presented. Genetic cell lines will be made available to those upon request. Correspondence should be addressed to EAH (hendr064@umn.edu).

Tables

Table 5.1. Results of immortalization experiments

Experiment Number	Number of Initial Cells	# Colonies after 2-weeks	# Immortal Clones	# Correctly Targeted	#ALT
1	1 x 10^6	117	1	1	1
2	1 x 10^6	128	2	2	2
3	1 x 10^6	143	13	13	13
PIGA-Neo Control	1 x 10^6	173	0	0	0

Figures

Figure 5.1

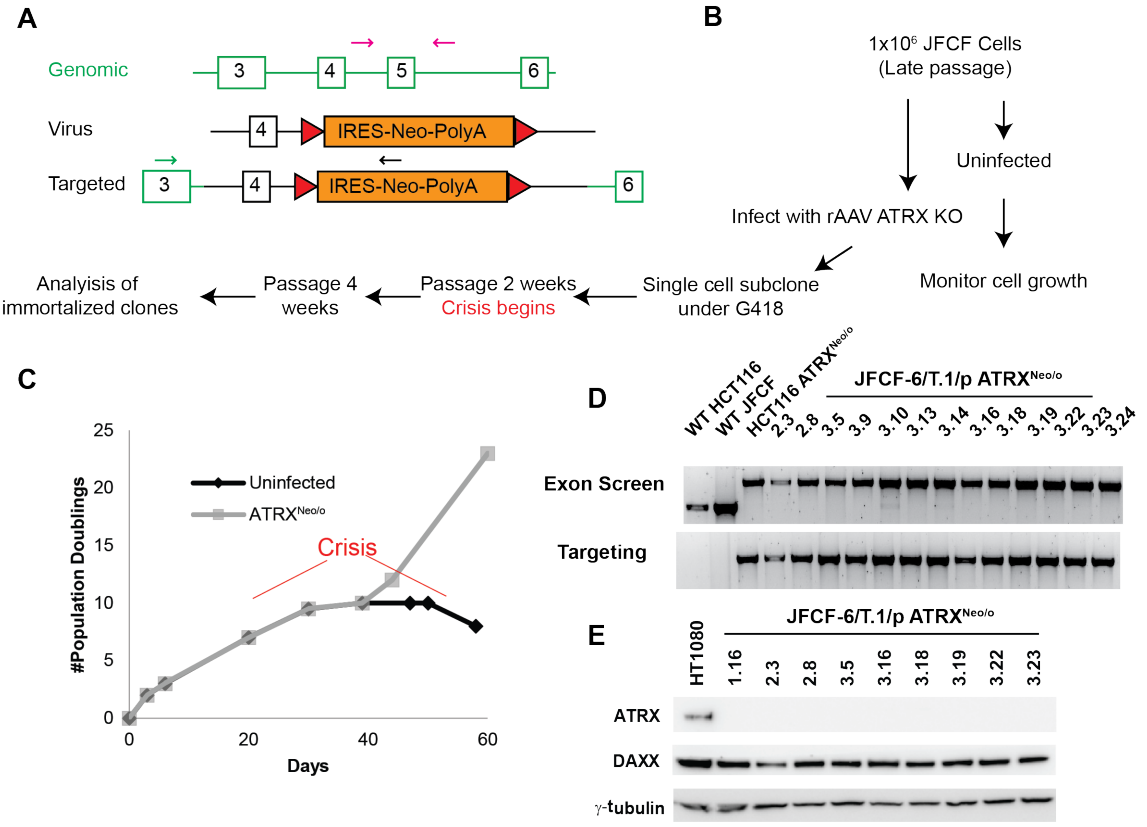


Figure 5.1. Genetic knockout of ATRX in pre-crisis cells correlates 100% with immortalization.

A) Schematic of the rAAV-ATRX-Neo gene targeting construct. Correct gene targeting is facilitated by HR-dependent gene replacement. Pink arrows correspond to "exon screen" PCR primers and those in black and green are "targeting" PCR primers. B) Schematic of the experimental setup. C) Growth chart monitoring population doublings over time. D) PCR confirming that all immortalized clones were correctly gene targeted by rAAV-ATRX-

Neo. E) Western blot analysis confirming the loss of ATRX protein expression in all the immortalized clones.

Figure 5.2

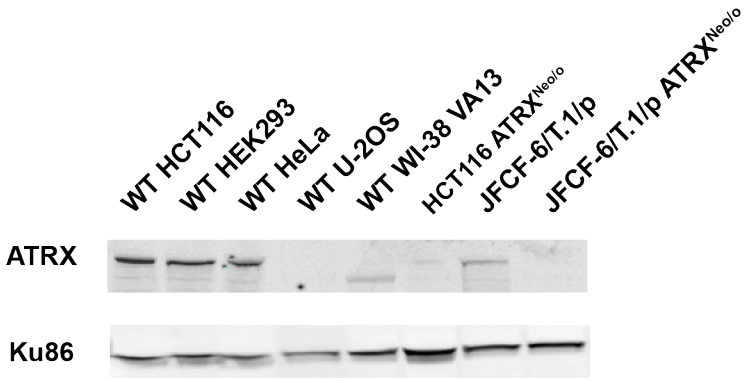


Figure 5.2. JFCF cells express ATRX protein.

Cell lysates from the indicated cell lines were collected and a subjected to a western blot for ATRX. Pre-crisis JFCF cells do express ATRX, which is abrogated by gene targeting.

Figure 5.3

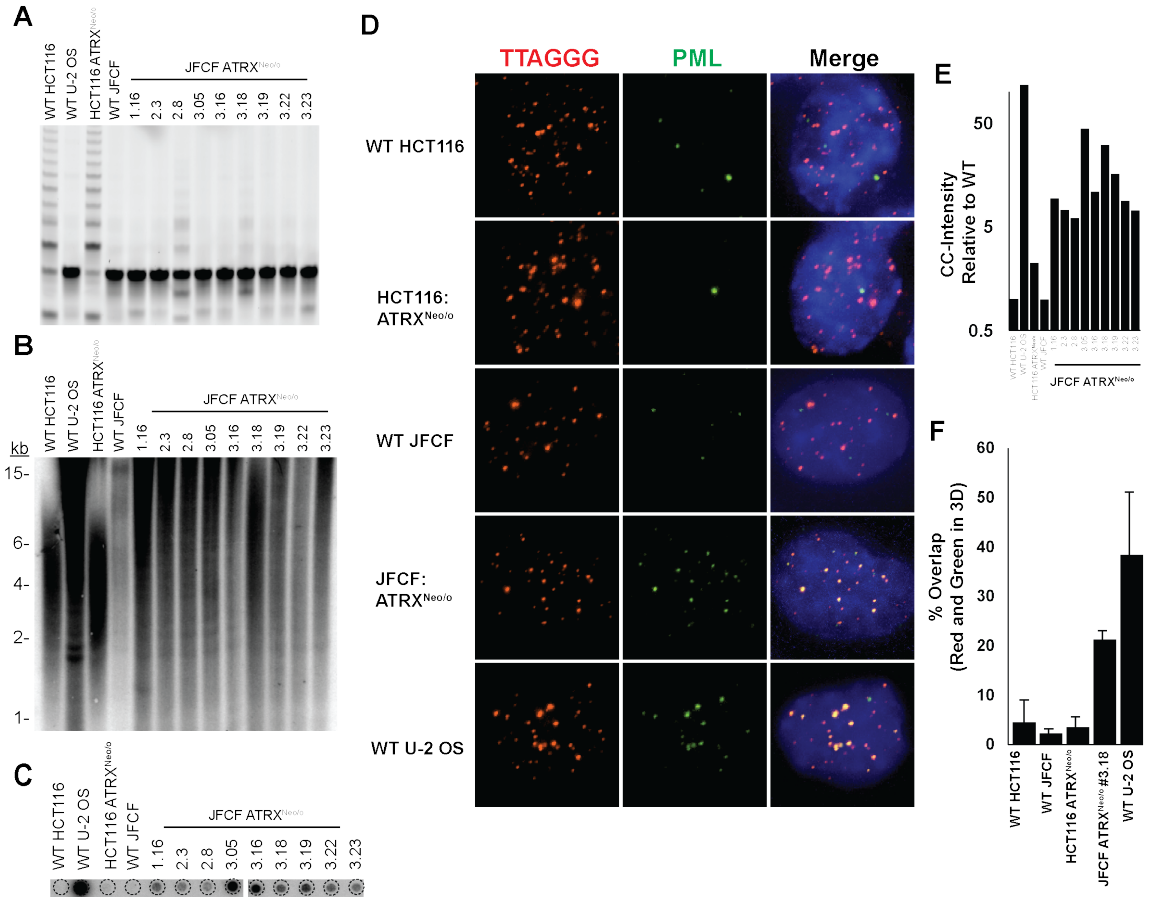


Figure 5.3. All cells that immortalize due to the loss of ATRX utilize ALT.

A) Representative TRAP assay depicting that all JFCF *ATRX*$^{Neo/o}$ clones have no telomerase expression. Wild-type HCT116 (WT HCT116) and U-2 OS were used as telomerase-positive and ALT-positive controls, respectively. B) A TRF length analysis shows that most of the JFCF *ATRX*$^{Neo/o}$ immortal clones have elongated telomeres. C) A C-circle assay demonstrates that all of the *ATRX*$^{Neo/o}$ clones are positive for the C-circle assay. D) Immunostaining for PML combined with a PNA telomere-FISH probe to analyze the incidence of APBs. There was a strong co-localization frequency between large nuclear PML bodies and large telomeric signals. E) Quantification of the signal intensity

of the C-circle assay shown in (C). F) Percent overlap of the PML- and telomere channels from (D) derived by computational 3D analyses.

Figure 5.4

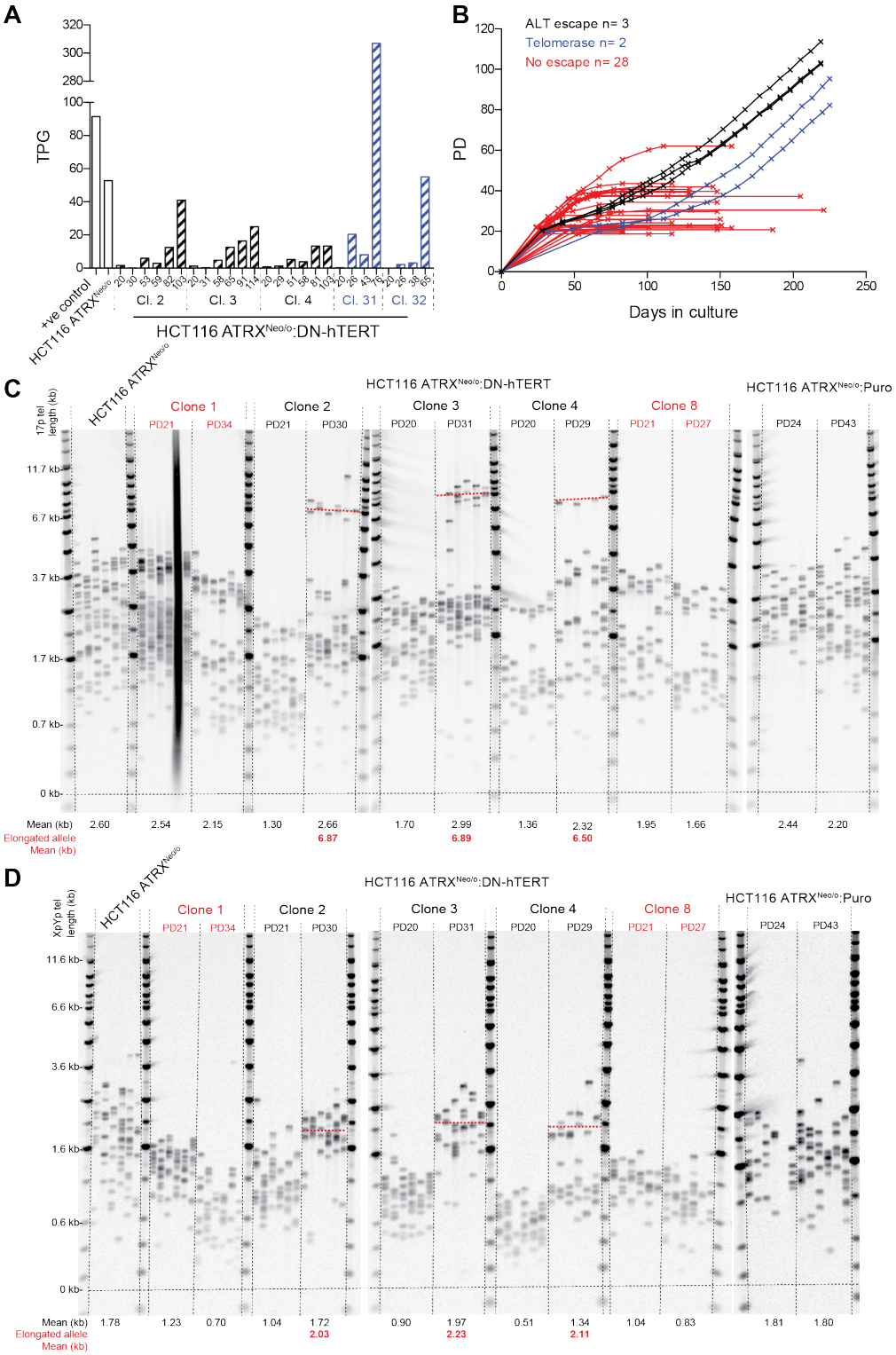

A
TPG
+ve control
HCT116 ATRX Neo/o
Cl. 2
Cl. 3
Cl. 4
Cl. 31
Cl. 32
HCT116 ATRX Neo/o:DN-hTERT
B
ALT escape n= 3
Telomerase n= 2
No escape n= 28
PD
Days in culture
C
HCT116 ATRX Neo/o
HCT116 ATRX Neo/o:DN-hTERT
HCT116 ATRX Neo/o:Puro
17p tel length (kb)
Clone 1
Clone 2
Clone 3
Clone 4
Clone 8
Mean (kb)
Elongated allele Mean (kb)
D
HCT116 ATRX Neo/o
HCT116 ATRX Neo/o:DN-hTERT
HCT116 ATRX Neo/o:Puro
XpYp tel length (kb)
Clone 1
Clone 2
Clone 3
Clone 4
Clone 8
Mean (kb)
Elongated allele Mean (kb)

Figure 5.4. Telomere crisis in ATRX-null HCT116 cells enables ALT-like activity.

Telomerase activity was abrogated following the expression of DN-hTERT in *ATRX*$^{Neo/o}$ HCT116 cells to induce telomere attrition over time. A) Quantification of telomerase activity plotted as total product generated (TPG). B) Population doublings (PDs) indicate that telomere crisis occurred approximately 20 PDs post-DN-hTERT expression. Five independent clones survived crisis (blue and black lines) and 28 clones did not (red). C-D) STELA analysis of the 17p and XpYp telomeres in *ATRX*$^{Neo/o}$ HCT116 cells expressing DN-hTERT or an empty vector control at PDs from the point of single cell cloning as indicated above. Clones 1 and 8 (in red) did not escape telomere crisis, and correspondingly exhibited no telomere elongation event, as did clones 2,3 and 4. Mean telomere is indicated below in black, telomere length of elongated telomeres is shown in red and highlighted with a dotted red line on the gel image.

Figure 5.5

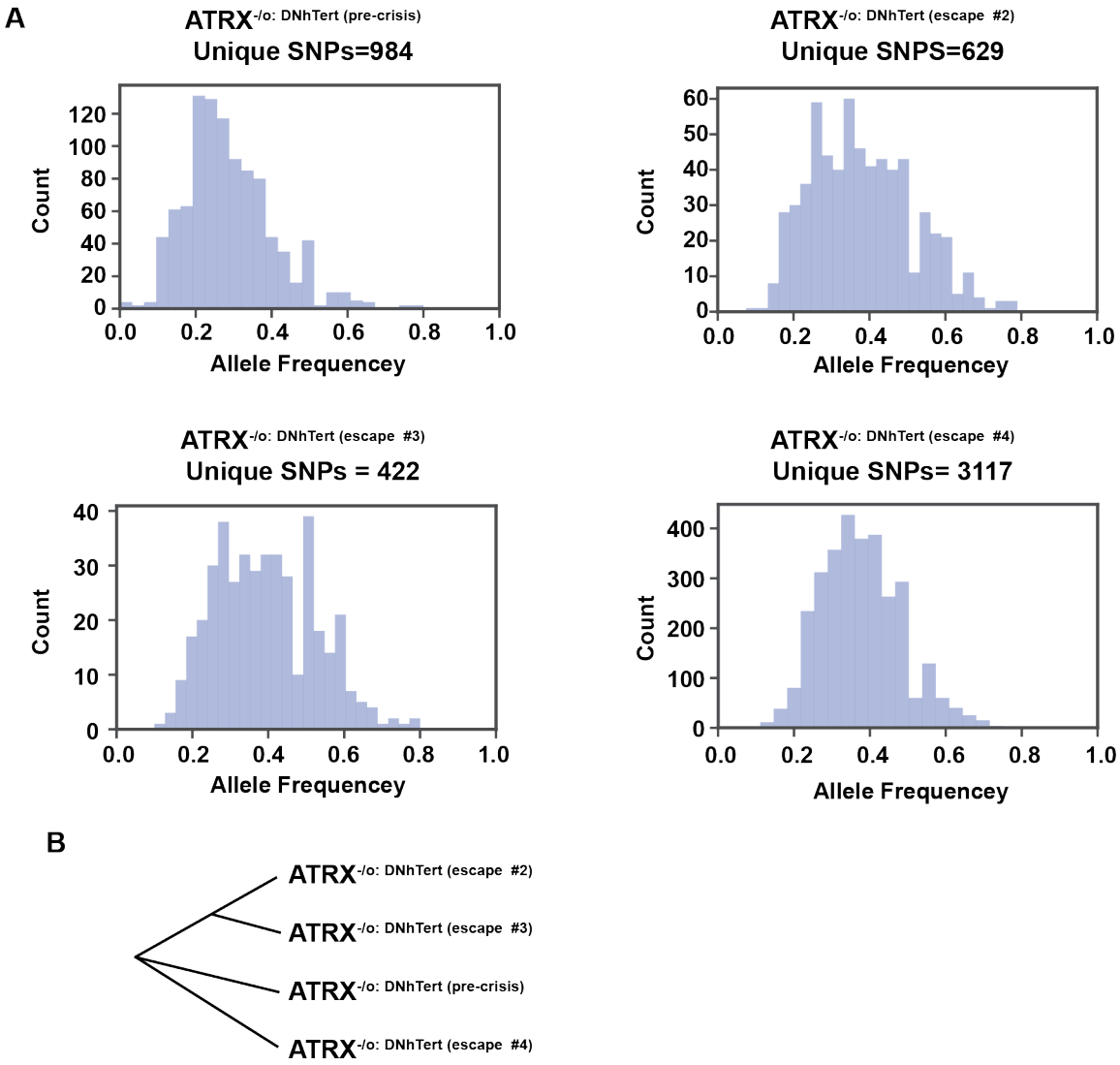


Figure 5.5. ALT-like clones that escaped crisis are genetically distinct.

A) Clones that escaped crisis were compared to the pre-crisis parental clone by NGS. Unique SNPs were unbiasedly identified, and plotted here. B) Phylogenetic analysis of the SNPs reveals that the clones post-crisis are genetically distinct.

Figure 5.6

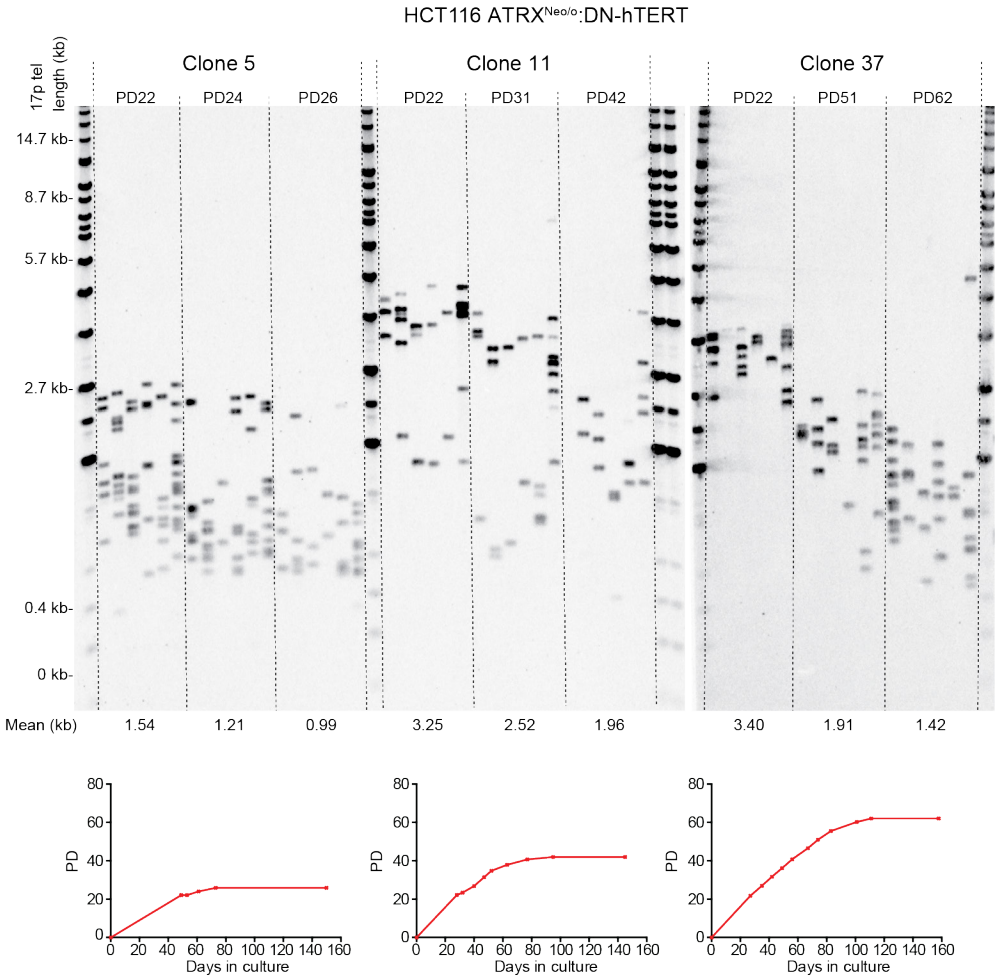


Figure 5.6. Clones that do not escape crisis are unable to maintain telomere length. STELA of the 17p telomere in serial passages of three clonal cultures of *ATRX*$^{Neo/o}$ HCT116 cells expressing DN-hTERT; all three clones failed to escape crisis. Clones and PD points are detailed above with mean telomere length and growth curves below.

Figure 5.7

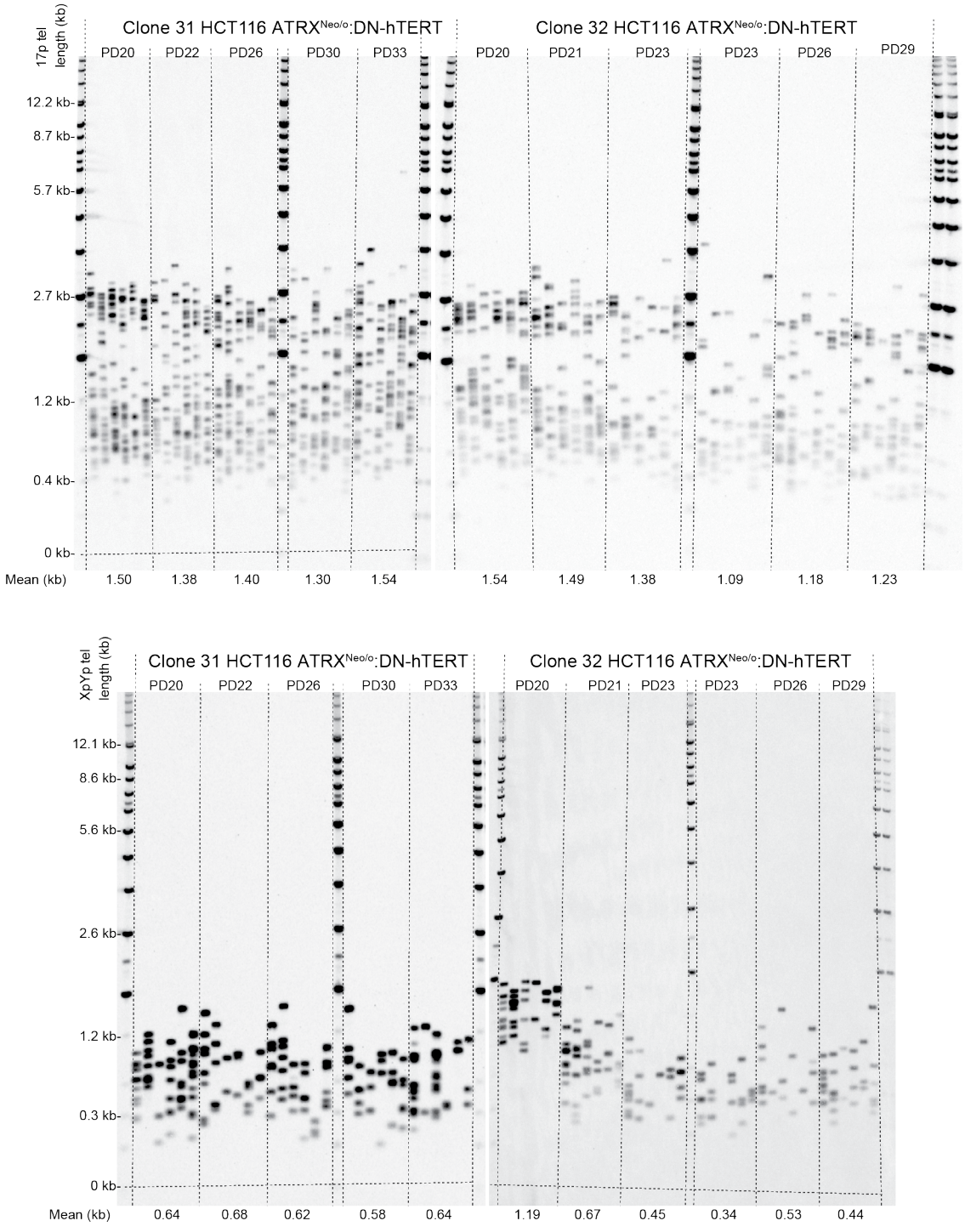

17p tel length (kb)
Clone 31 HCT116 ATRX^Neo/o:DN-hTERT
PD20
PD22
PD26
PD30
PD33
Clone 32 HCT116 ATRX^Neo/o:DN-hTERT
PD20
PD21
PD23
PD23
PD26
PD29
12.2 kb-
8.7 kb-
5.7 kb-
2.7 kb-
1.2 kb-
0.4 kb-
0 kb-
Mean (kb)
1.50
1.38
1.40
1.30
1.54
1.54
1.49
1.38
1.09
1.18
1.23
XpYp tel length (kb)
Clone 31 HCT116 ATRX^Neo/o:DN-hTERT
PD20
PD22
PD26
PD30
PD33
Clone 32 HCT116 ATRX^Neo/o:DN-hTERT
PD20
PD21
PD23
PD23
PD26
PD29
12.1 kb-
8.6 kb-
5.6 kb-
2.6 kb-
1.2 kb-
0.3 kb-
0 kb-
Mean (kb)
0.64
0.68
0.62
0.58
0.64
1.19
0.67
0.45
0.34
0.53
0.44

Figure 5.7. Two clones that escaped crisis without allele specific telomeric elongation.

STELA of the 17p telomere in serial passages of a long-term culture of *ATRX*$^{Neo/o}$ HCT116 cells expressing DN-hTERT. PD points are detailed above, with overall mean telomere length detailed below, together with the long or short allele specific means shown in red and green. The changes in telomere length (Δ bp/PD) of the two alleles are detailed.

Figure 5.8

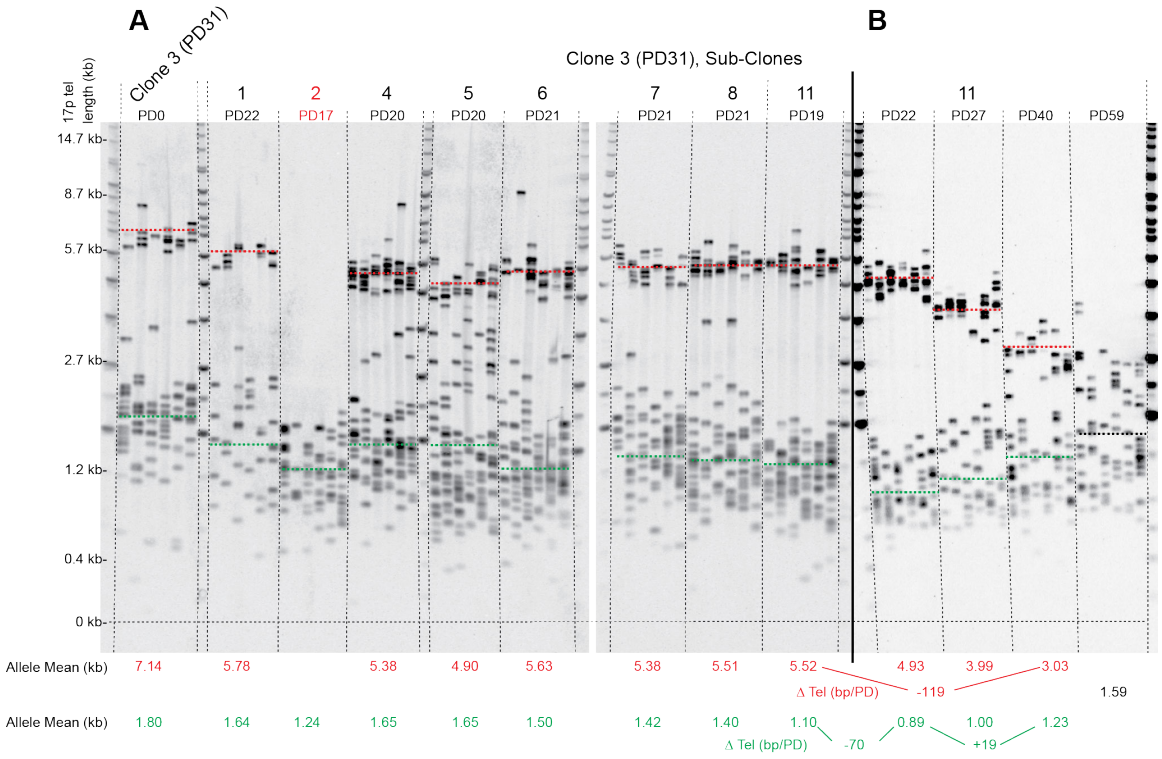


Figure 5.8. Survival is dependent on multiple, end-specific, telomere elongations in ATRX-null cells undergoing telomere crisis.

Clone 3 of *ATRX*$^{Neo/o}$ HCT116 cells expressing DN-hTERT that survived crisis was single cell sub-cloned at PD31, and analyzed by STELA. A) Only clones that exhibited the specific 17p elongation events were able to survive, with one that did not (subclone 2, in red) dying at PD17. Sub-clone numbers and PD points from the point of single cell sub-cloning are detailed above, with the longer and shorter allelic telomere length distributions indicated in red and green, respectively. B) Sub-clone 11 was passaged up to PD59 and telomere length monitored, the changes in telomere length (Δ bp/PD) of the two alleles are detailed.

Figure 5.9

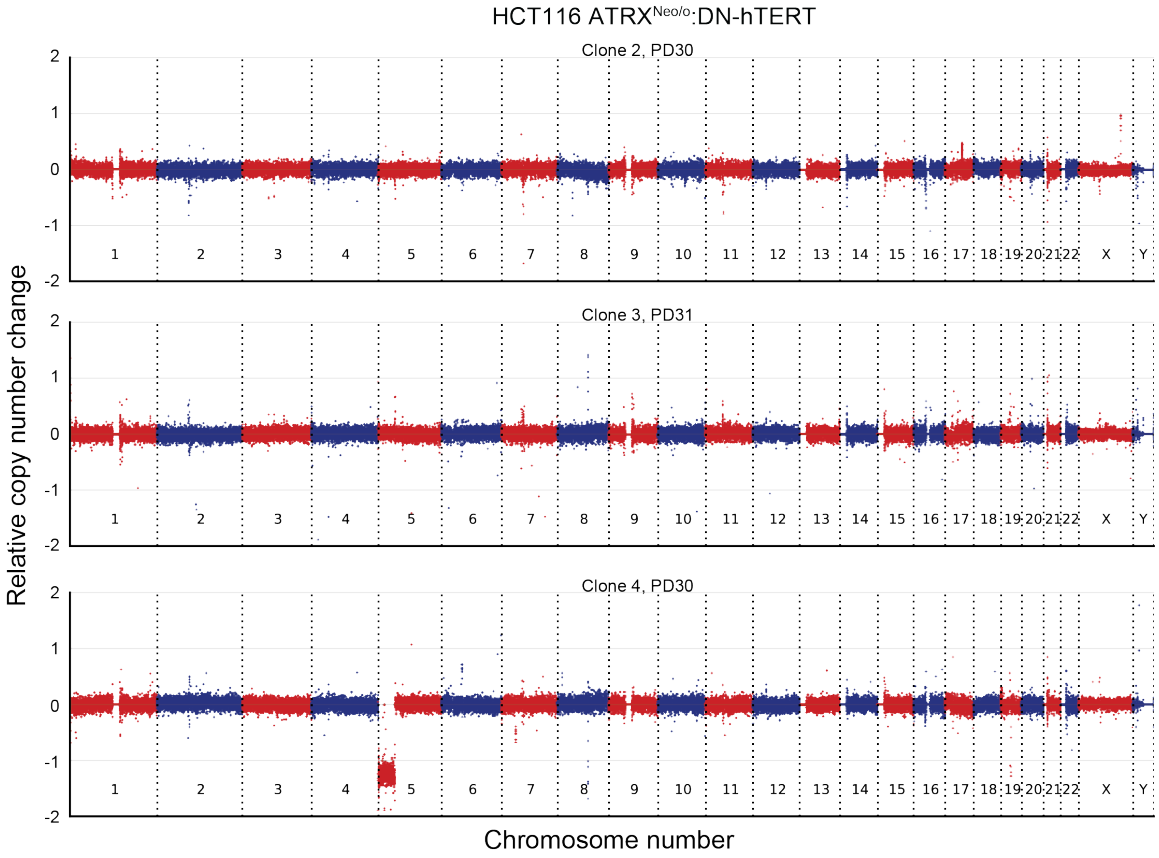


Figure 5.9. Telomeric elongation in the absence of ATRX in cells undergoing a telomeric crisis is not accompanied by large-scale genomic rearrangements.

Relative copy number change profiles from whole genome sequencing data derived from three *ATRX*$^{Neo/o}$ HCT116 cells expressing DN-hTERT that escaped crisis with telomeric elongation.

Figure 5.10

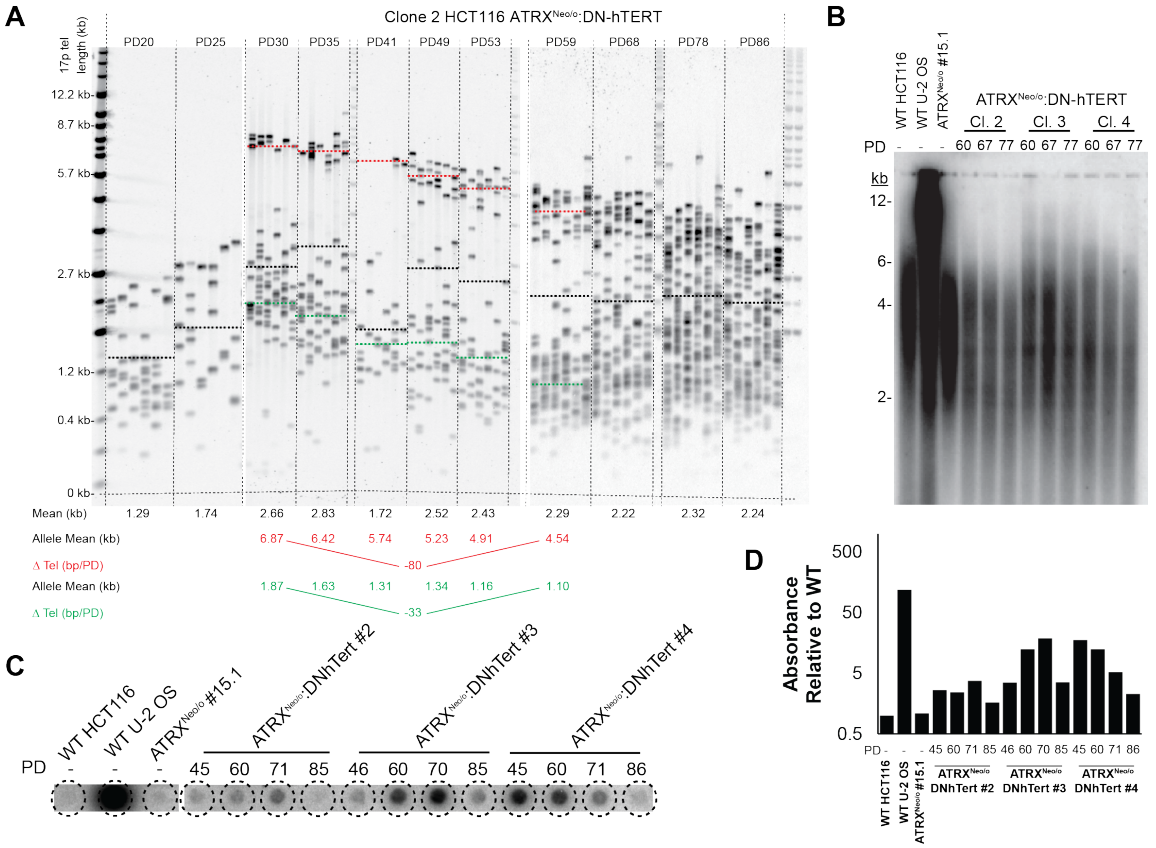


Figure 5.10. The absence of ATRX is not sufficient to maintain an ALT phenotype. A) STELA of the 17p telomere in serial passages of a long-term clonal cultures of *ATRX*$^{Neo/o}$ HCT116 cells expressing DN-hTERT. Following the initial elongation event observed at PD 30, both telomeric alleles exhibited telomere erosion with ongoing cell division. PDs points are detailed above, with overall mean telomere length detailed below, together with the long or short allele specific means shown in red and green. The changes in telomere length (Δ bp/PD) of the two alleles are detailed. B) A TRF depicting telomere lengths at late PDs after crisis, which depicts heterogeneous telomeres that shorten and compact over time. C) C-circle analysis of clones throughout the aging of the HCT116

DN-hTERT-expressing cells, depicting intermittent ALT activity. D) Quantification of the absorbance depicted in (C) relative to WT in log-scale.

Figure 5.11

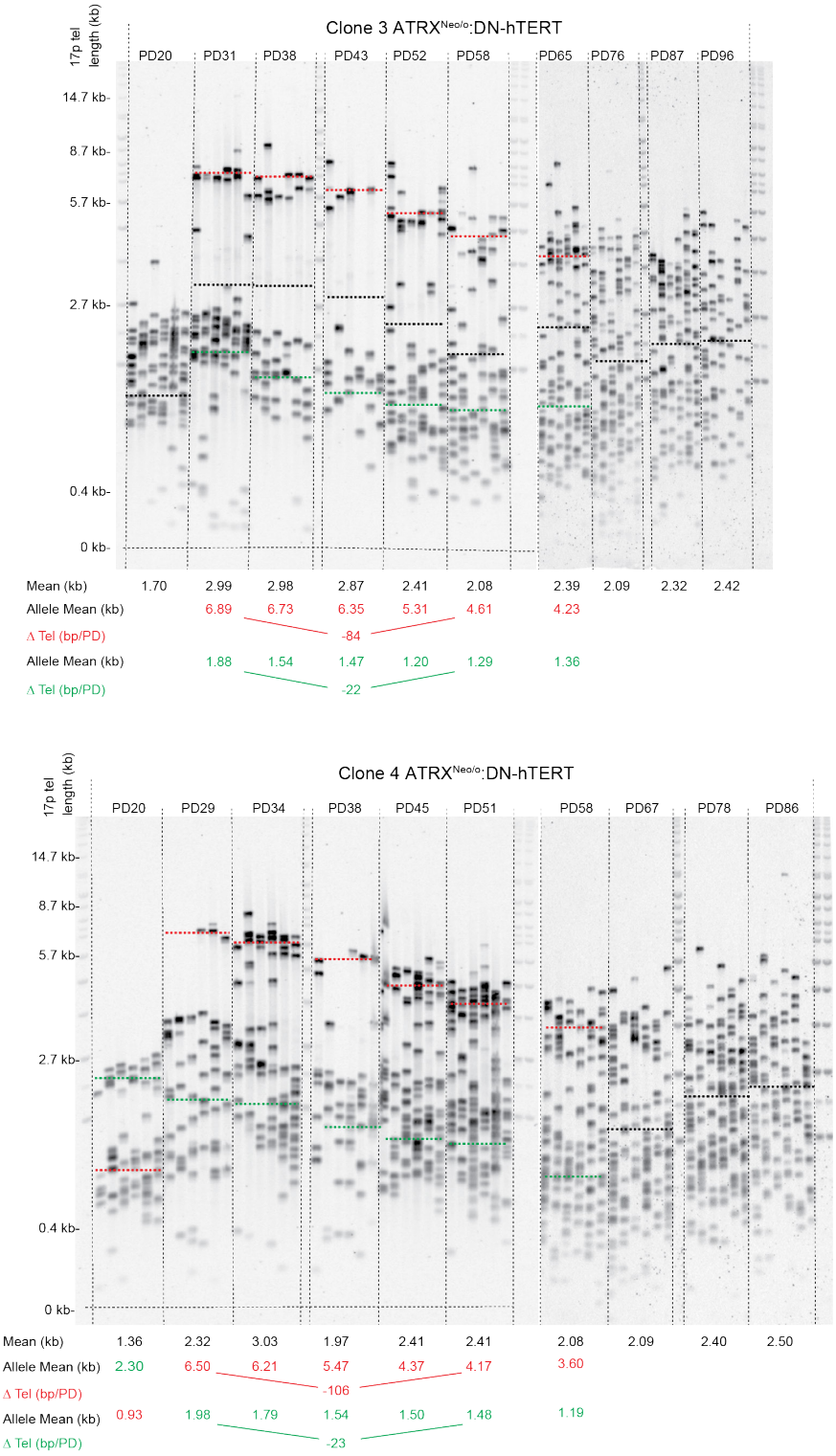

Clone 3 ATRX^Neo/o:DN-hTERT
17p tel length (kb)
PD20 PD31 PD38 PD43 PD52 PD58 PD65 PD76 PD87 PD96
14.7 kb-
8.7 kb-
5.7 kb-
2.7 kb-
0.4 kb-
0 kb-
Mean (kb) 1.70 2.99 2.98 2.87 2.41 2.08 2.39 2.09 2.32 2.42
Allele Mean (kb) 6.89 6.73 6.35 5.31 4.61 4.23
Δ Tel (bp/PD) -84
Allele Mean (kb) 1.88 1.54 1.47 1.20 1.29 1.36
Δ Tel (bp/PD) -22
Clone 4 ATRX^Neo/o:DN-hTERT
17p tel length (kb)
PD20 PD29 PD34 PD38 PD45 PD51 PD58 PD67 PD78 PD86
14.7 kb-
8.7 kb-
5.7 kb-
2.7 kb-
0.4 kb-
0 kb-
Mean (kb) 1.36 2.32 3.03 1.97 2.41 2.41 2.08 2.09 2.40 2.50
Allele Mean (kb) 2.30 6.50 6.21 5.47 4.37 4.17 3.60
Δ Tel (bp/PD) -106
Allele Mean (kb) 0.93 1.98 1.79 1.54 1.50 1.48 1.19
Δ Tel (bp/PD) -23

Figure 5.11. The absence of ATRX is not sufficient to maintain an ALT phenotype. STELA of the 17p telomere in in serial passages of two long term clonal cultures of *ATRX*$^{Neo/o}$ HCT116 cells expressing DN-hTERT. Following the initial elongation event observed at PD 30, both telomeric alleles exhibit telomere erosion with ongoing cell division. PD points are detailed above, with overall mean telomere length detailed below, together with the long or short allele specific means shown in red and green. The changes in telomere length (Δ bp/PD) of the two alleles are detailed.

Chapter 6

Discussion and Future Directions

Implications of my work

The work I have described here is focused on both the details of — and the connections between — DNA DSB break repair, telomere maintenance, and cellular immortalization. First, in **Chapter 2**, I showed how the subpathways of NHEJ are genetically regulated and identified the existence of a Ku-independent, LIG4-dependent NHEJ reaction. As discussed, the function of telomeres is to repress the recognition of the end of a linear chromosome as a DNA DSB. Ironically should that occur, which DSB repair pathway mistakes the end of the chromosome for a DSB dramatically influences whether the cell will die, or whether it will have a chance to survive and escape the ensuing telomere crisis and potentially immortalize. In **Chapter 3**, I showed that PARP1 is an important regulator of telomere homeostasis, and its mutation is causative of DNA damage at telomeres and that it results in telomere shortening. In **Chapters 4 and 5**, I described how the genetic inactivation of *ATRX*, which is strongly associated with ALT-positive cancers, is not a *bona fide* repressor of ALT, but that its mutation does dramatically allow for ALT to occur, presumably due to ALT-like telomere recombinations. Taken together, in the work presented throughout this dissertation, I have attempted to gain new insight into the regulation of DNA DSB repair, how some of these repair proteins regulate normal telomere homeostasis, and how ALT immortalization is functionally repressed by ATRX.

Chapter 2 is focused on the regulation of NHEJ, and how C-NHEJ influences A-NHEJ. I showed that while *LIG3*-null cells are largely aphenotypic when it comes to DNA repair and survival, in Ku-deficient cells, LIG3 is primarily responsible for A-NHEJ. This likely underlies the many observations that A-NHEJ, while rare, is nonetheless consistently detectable (Simsek and Jasin 2010). Independently, other members of our laboratory

showed that *LIG3* was essential for telomere immortalization, whereas *LIG4* rather mediated lethal interchromosomal fusions (Oh et al. 2014). Thus, one implication of my work is that the *Lig3*-dependent telomere fusions are repressed by Ku. This implication fits nicely with the telomere phenotypes observed in the Ku-deficient cells, which appeared to consistently undergo intrachromosomal fusions (Wang et al. 2009). Furthermore, the requirement for a rare DNA repair event for cellular immortalization is consistent with the low incidence of cells in telomere crisis that survive and immortalize.

Another implication of my work in **Chapter 2** is that there is a requirement for *LIG1* in some form of end-joining, which arises from the observation that while all of the Ku-repressible A-NHEJ is *LIG3*-dependent, there is always some level of A-NHEJ detectable in the absence of both proteins (Figs 2.9 and 2.10). Notably, deletion of *LIG4* did not abolish A-NHEJ levels in either WT or Ku-deficient cells, thus precluding any role for *LIG4* in A-NHEJ (Fig 2.10). The only remaining candidate DNA ligase is *LIG1*, which presumably must be functioning to enable A-NHEJ. There are two possible models to explain how *LIG1* is participating in A-NHEJ; either *LIG1* is functionally redundant with *LIG3*, or there are at least two subpathways of A-NHEJ present in human cells. Support for the functional redundancy between *LIG3* and *LIG1* comes from the evolutionary history of *LIG3*. Many organisms, including the model organism, *Saccharomyces cerevisiae*, do not have a third DNA ligase, and in such instances, this role is filled by DNA *LIG1* (Simsek and Jasin 2011). As the essential role for *LIG3* is in the mitochondria, it appears that *LIG3* is required during mitochondrial DNA replication for sealing Okazaki fragments, thus implicating a functional redundancy between it and the primary role of *LIG1* in the nucleus of eukaryotes (Simsek et al. 2011). Furthermore, the complementation of *LIG3*-null cells with a mitochondrially-directed *LIG1* can completely rescue *LIG3*'s essential function

(Simsek et al. 2011). Our laboratory has recently constructed *LIG1*$^{F/-}$ conditionally-null cells. As has been previously observed in most, but not all species, *LIG1* is essential in humans. Thus, while I have shown in **Chapter 2** that *LIG3* may be partially redundant with *LIG1*, the converse is not true; *i.e. LIG3* cannot compensate for the absence of *LIG1*.

With respect to the two possibilities of *LIG3* and *LIG1* either being functionally redundant or being discrete subpathways of A-NHEJ, I favor the later argument. Models based off the earlier work in budding yeast have predicted this idea, which describe two subpathways of A-NHEJ (Decottignies 2013). These models predict a *LIG3*-dependent, HDR-breakdown pathway, consistent with previous observations that the upstream HDR genes impact A-NHEJ (Rass et al. 2009). The other pathway is proposed to be a polymerase-dependent as well as replication-dependent pathway, which would require specific polymerases and *LIG1*.

This *LIG1*-dependent A-NHEJ pathway is intriguing as recent work in the context of cellular immortalization has shown that telomeres can undergo intra-chromosomal fusions in the absence of both *LIG3* and *LIG4*, implicating a role for *LIG1* in such events (Liddiard et al. 2016). This was surprising, as a requirement for *LIG3* in cellular immortalization was presumed to be due to *LIG3* being responsible for a discrete end-joining pathway which permitted survival (Jones et al. 2014); however, *LIG1* appears to be able to perform such repair events as well. These two observations beg the question, “what is the essential role for *LIG3* in cellular immortalization”, especially if *LIG1* is redundant with *LIG3*? One simple answer to this question might be that if *LIG3* is epistatic to *LIG1*, then the genetic requirement for *LIG3* may be to prevent *LIG1*-dependent A-NHEJ. However, other explanations are also plausible and should be addressed in future studies. This emerging

subpathway of A-NHEJ may be of significant clinical interest, as it likely influences the immortalization of many different cancers.

In **Chapter 3** I characterized the role of *PARP1* in human cells, which was initiated because of previous implications of *PARP1* in A-NHEJ (Wang et al. 2006). However, unlike in the case of *LIG3*, I was unable to detect any significant contribution of PARP1 to A-NHEJ. I attribute the differences to likely be a combination of the organisms in which previous data were collected and the difference between chemical inhibition of PARP1 and the genetic ablation of *PARP1*. I have discussed these differences in both **Chapter 1** and **Chapter 3** in detail. Yet, the more surprising discovery I present in **Chapter 3** was a strong role of *PARP1* in maintaining telomere homeostasis.

Telomeres are notoriously difficult regions of the genome to replicate, and are considered genomic fragile sites (Sfeir *et al.* 2009). They are prone to the formation of G-quadruplexes, are inherently repetitive, and are unlikely to be rescued by dormant origins in the case of replication stalling (Gilson and Geli 2007). Preferential recruitment of telomere accessory factors such as WRN and BLM by Shelterin are critical for telomere stability as the telomeres are replicated (Ye *et al.* 2010). Notably, mutations in *WRN* are causative of Werner's syndrome, which is a disease of premature aging (Martin 1978), a classic phenotype for telomere dysfunction diseases. Thus, WRN likely functions to enable proper telomere replication, and its absence is causative of accelerated ageing.

Importantly, PARP1 physically interacts with WRN (von Kobbe *et al.* 2003), which implicates them in a similar or common pathway. PARP1 also travels with the replication fork and is responsible for DNA damage signaling specifically related to single-strand breaks (Min et al. 2013). Thus, it seems likely (albeit as of yet unproven) that PARP1 and WRN may cooperate to help with stalled replication forks. While compensatory

mechanisms exist for dealing with stalled forks in genomic regions, fragile sites are more prone to DNA damage and fork stalling when they are located in telomeric regions. It is for this reason that I suggest that PARP1 (and probably WRN) is preferentially recruited to telomeres to ensure the fidelity of their replication.

In **Chapter 3**, I present evidence that *PARP1*-null cells have short telomeres and that spontaneous DNA damage markers are enriched at these telomeres. It is important to note, however, that the absence of PARP1 may only be tolerated in a human cancer cell that is over-expressing telomerase. Notably, the asymmetry of the gene targeting I observed in the generation of *PARP1*$^{-/-}$ cells suggested that there is a strong evolutionary selection to retain at least one allele of *PARP1* (Table 3.1). This observation may also explain why there are no known inherited *PARP1* genetic diseases in humans. While the role of *PARP1* in DNA repair may have redundant components, its ancillary role in enabling sufficiently long and stable telomeres may have been more strongly selected for. This is especially intriguing, as *PARP1*-null mice are both viable and fertile (Wang et al. 1995). Taken together, my results suggest that there may be a uniquely essential role for *PARP1* in the telomere maintenance of healthy human cells. The use of CRISPR/Cas9 to try editing the PARP1 gene in normal (non-immortalized, non-transformed) fibroblasts would be one way to experimentally address this hypothesis.

In **Chapter 4**, I describe the foundation of a large series of work regarding the role of *ATRX* in ALT. Given the incredibly strong association between ALT and *ATRX* mutations, it was surprising that the genetic mutation of *ATRX* in telomerase-positive cells did not convert them to ALT. This was especially true given that mutation of a different H3.3 chaperone, *ASF1*, was sufficient to induce ALT activity in telomerase-positive cells (O'Sullivan et al. 2014). These results are seemingly incongruent, as *ATRX* mutations are

very frequent in ALT cancers whereas *ASF1* mutations are not. Nonetheless, the relatively high frequency of *ATRX* mutations in telomerase-positive cancers supports my finding that the absence of ATRX appears to be necessary for the initiation but not the maintenance of ALT, as there are also many known cancers with *ATRX* mutations that are not ALT (Fishbein *et al.* 2015). This complex genotype to phenotype correlation is most likely a reflection of the role of ATRX in regulating both DNA replication and DNA repair (Leung et al. 2013), as it has been well established that ATRX facilitates chromatin remodeling in regions other than just the telomere (Goldberg et al. 2010). Thus, the loss of ATRX may impact multiple regulatory processes required for the Darwinian selection occurring in the oncogenic development of a malignancy. Since some of these other regulatory processes have built in redundancies and some do not, the requirement for the loss of ATRX is not consistent from cell type to cell type and from tumor to tumor. This may also explain why in ALT cancers, mutation of ATRX is far more common than mutation of DAXX (whose only known function is that of a histone H3.3 chaperone), although they function in the same telomere-specific H3.3 histone deposition pathway.

I thus next tried to address the extent to which *ATRX* regulates the genesis of ALT in **Chapter 5**. My primary finding was that the inactivation of *ATRX* significantly enhances the frequency of ALT by between 4 to 5 orders of magnitude. This observation, in and of itself, probably explains the frequency of *ATRX* mutations in ALT. From a molecular perspective, my work also provides insight into what role ATRX is likely normally playing in preventing ALT from initiating, which is that it represses the recombination of telomeres during crisis. My work also suggests that the telomeric recombinations may be impacted by *cis*-acting DNA regulatory elements, likely specific to each sub-telomere, as the recombinations in multiple independent clones resulted in elongation to a specific length.

One fascinating explanation for these site-specific recombinations could be that the incorporation of H3.3 by ATRX is tightly controlled by sequence specific elements. One report found ATRX-dependent H3.3 incorporation into regions of the genome in areas as little as ~100 bp long (Levy *et al.* 2015). These areas are predicted to be sites prone to the formation of G-quadruplex DNA, structures that are also common in telomeres (Smith and Feigon 1992). Thus, the site-specific recombinations seen in the *ATRX*$^{Neo/o}$ clones during their escape from crisis may be correlated with regions normally occupied by H3.3.

Future Directions

In the remaining text, I will discuss outstanding questions relevant to my work and propose experiments that could address these unknown issues.

1) What is the mechanistic contribution of PARP1 to telomere maintenance?

My work implicates PARP1 as a suppressor of stochastic replication-induced DNA damage at a telomere. However, it is unclear whether this role is simply due to PARP1's ubiquitous role in resolving SSBs and damage at a replication fork, or if there exist discrete (identifiable perhaps by separation-of-function mutations) mechanisms. Thus, I would propose to determine if PARP1's recruitment to the telomere via its interaction with TRF2 is required to prevent the damage that occurs in the telomeres during its absence. Notably, although this interaction has been biochemically confirmed (Gomez et al. 2006), the functional consequence of this interaction remains speculative. PARP1 contains a TRF2-docking motif 737-Y-T-L-I-P-741, and the mutation of this motif in other proteins abrogates their telomere localization (Chen et al. 2008). Thus, I would propose to

complement the PARP1-null cells with a mutant *PARP1*$^{\Delta TRF2}$ cDNA and propose the following experiments.

Sub-aim 1: Determine the functional consequence of PARP1's interaction with TRF2

In *PARP1*-null cells complemented with either, WT, *PARP1*$^{\Delta TRF2}$ or empty vector, I would address PARP1 localization to the telomere by a ChIP experiment. Telomere ChIPs are relative easy to perform because they can utilize a well-optimized telomeric DNA probe for detection by Southern blotting. Cell lysates from each of the genotypes would be collected, and the proteins would be crosslinked to the DNA with 4% formaldehyde. I would then IP PARP1, dot blot the lysate onto a membrane and probe it using a γP-32 radiolabeled 5'-$(CCCTAA)_3$-3' oligonucleotide. Such techniques have been used to monitor other specific protein:telomeric DNA binding interactions (Sfeir and de Lange 2012). The telo-ChIP approach is especially quantitative as it does not require any additional PCR amplification. I would hypothesize that PARP1's recruitment to the telomere would be dramatically reduced in the *PARP1*$^{\Delta TRF2}$ cells. Nonetheless, PARP1 is a SSB binding protein, and if such a break does occur anywhere in the genome, it will likely still be able to access it (the TRF2 interaction domain does not overlap with PARP1's DNA binding domain) (Satoh and Lindahl 1992). To understand if PARP1's interaction with TRF2 is required for DNA damage sensing at the telomere, I would treat the cells with 1 Gy of ionizing radiation, which while known to produce ~10 to 20 DSBs per cell, is estimated to cause ~30X the amount of SSBs/cell than DSBs/cell (Blocher and Pohlit 1982). I hypothesize that while PARP1 usually binds to DNA breaks throughout the genome, at the telomere it is occluded from binding to such breaks without its proper

interaction with TRF2, as Shelterin is well documented to prohibit DNA repair enzymes from accessing the telomere to initiate DNA repair (de Lange 2010). Thus, I would repeat the telomere-ChIP assay with the aforementioned panel of cell lines, to determine if PARP1's binding to the telomere can be enhanced by exogenous damage. The relative amount of DNA damage +/- radiation could be confirmed by immunofluoresence for the DNA DSB markers 53BP1 and γH2AX.

Lastly, I have suggested that PARP1's role at the telomere is likely to prevent replication-induced DNA damage, which if unrepaired, would be causative of telomere shortening. My hypothesis is that PARP1's prevention of replication-induced DNA damage requires its interaction with TRF2 and localization to the telomere. To evaluate if PARP1 is contributing to the fidelity of telomeric replication, I would perform telomere molecular combing. This technique combines the molecular combing of DNA, which linearly stretches genomic DNA on a microscope slide, with telomere FISH (Suram *et al.* 2012). Molecular combing is a technique that stretches out a chromosome as one single molecule on a microscope slide, which enables the examination of the molecular dynamics of various DNA loci on the same chromosome (Lebofsky and Bensimon 2003). One application of molecular combing is the visualization of DNA replication orientation and progression along a chromosome by utilizing pulse-chase experiments with nucleotide analogs that can be detected by fluorescently conjugated antibodies. Following molecular combing, to visualize the telomere, telomeric FISH is performed to label the telomere (Suram et al. 2012). Thus, this hybrid assay enables the examination of whether telomeres become fully replicated. To determine if PARP1 is facilitating proper telomeric replication, I would perform this assay on each of the *PARP1*-null cells complemented with either WT, *PARP1*$^{\Delta TRF2}$, or empty vector. Furthermore, to enhance the frequency of these

replication-induced damage events, I would treat each of these cells with or without aphidicolin, which is a DNA polymerase inhibitor and should thus increase the number of replication-induced DNA damage events per cell. Based on the observation that I detected ~2 to 3 TIFs per cell in untreated *PARP1*$^{-/-}$ cells (Fig 3.9), I would predict that the *PARP1*-null cells exhibit some detectable un-replicated telomeres even if untreated. I would further predict that these cells would be especially sensitive to aphidicolin treatment. Additionally, I would expect that the *PARP1*$^{\Delta TRF2}$ cells would phenocopy the *PARP1*-null cells for their capacity for telomere replication. One final interesting examination would be the difference between the genome-wide effects on replication fidelity of the *PARP1*$^{-/-}$ *and* *PARP1*$^{\Delta TRF2}$ cells, where I posit that if there is a true separation-of-function between these two mutants, then the *PARP1*$^{\Delta TRF2}$ cells would be proficient for fork stabilization throughout the genome.

Sub-aim2: Determine if telomere maintenance is PARP1's essential function in human cells

The gene targeting experiments in **Chapter 3** revealed that human cancer cells prefer to maintain and express *PARP1*, as only 3/72 targeted cells were correctly targeted (Table 3.1). During these gene targeting experiments there was no evidence of any haploinsufficiency of *PARP1*$^{+/-}$ cells, thus it may be possible that human heterozygotes are viable, but that homozygous nulls are embryonically lethal — in any event a human PARP1-null patient has never been described. This prediction is in stark contrast to the observation that *PARP1*-null mice are viable and fertile (de Murcia et al. 1997). This difference might be attributable to the fact that laboratory mice have many notable

differences compared to humans, including: having longer telomeres, expressing telomerase in most of their somatic cells, and being inherently resistant to telomere defects (Martinez and Blasco 2010). *PARP1*'s absence from a list of very similar DNA repair genes such as *ATM*, *ATR*, *NBS1*, *XPC*, and others, is notable, as the mutation of all of these other genes are each causative of uniquely corresponding human diseases (Negrini et al. 2010). My hypothetical explanation for this is that if *PARP1* is essential for proper telomere length maintenance (even in the presence of telomerase, as my work suggests), the loss of *PARP1* may cause developmental failure in humans, as human stem cells are especially sensitive to telomeric stress (Wong et al. 2010). In order to test this hypothesis, I propose to create conditionally-null $PARP1^{F/F}$ iPS cells using CRISPR/Cas9 gene editing. I would then subject these cells to differentiation assays as described (Weltner *et al.* 2012), either in the presence or absence of *PARP1*. I would predict that in the absence of *PARP1*, the proliferative capacity of these cells would be dramatically reduced. Notably, healthy stem cells would be more sensitive to DNA damage at telomeres than cancer cells, as they are not transformed and thus responsive to senescent/apoptotic signals (Ju and Lenhard Rudolph 2008). Thus, in these cells I would propose to monitor the kinetics of DNA damage at telomeres, either by immunofluorescence for TIFs, or by a telomere-ChIP for proteins such as 53BP1 or γH2AX. Additionally, in PARP1-null stem cells, I would expect to see evidence of cellular senescence, detectable using Western blot analyses for p21, a common senescence marker (el-Deiry et al. 1994). Moreover, I hypothesize that these anticipated growth and differentiation defects in $PARP1^{-/-}$ stem cells could be rescued genetically, via a compensatory mutation of *tp53*. Similar approaches in mice enable healthy MEFs to grow in the absence of fully functional telomeres (Celli and de Lange 2005). If the predicted

phenotypes were indeed observed, it would serve as further evidence that PARP1 is preventing a telomere shortening-cellular senescence program. In total, these experiments would determine if PARP1's role in preventing telomeric damage in healthy human cells is essential for life.

2) What additional factors are required for ALT-immortalization?

In **Chapter 5** I showed that mutation of ATRX in HCT116 is sufficient to enable critically shortened telomeres to recombine and escape crisis, in an ALT-like manner. This finding implicates a distinct role of ATRX in repressing telomere recombinations and some other factor(s) which promotes/enables the ALT-state. Hence it is critical to identify that second regulatory element(s) that is responsible for maintaining the ALT state. Here, I propose a genetic screen to identify such factors in human cells. Additionally, because *tp53* is very commonly mutated or inactivated in ALT tumors, I would seek to clarify if its mutation or misregulation is required and/or sufficient for the perpetuation of the ALT state.

Sub-aim1: Identify the factors responsible for perpetuating the ALT state

Recent work has reported that genetic mutation of *TERC* is an excellent tool to generate ALT cells (Min et al. 2017). This study utilized a fibroblast cell line, SW39, which is similar to JFCF and which can spontaneously immortalize using either ALT or telomerase. Given that in **Chapter 5** I could detect ALT-like activity, but not the continuance of an ALT-state, I suggest that while HCT116 is repressive of ALT, it retains the molecular capacity to produce it. Thus, HCT116 is an excellent cellular model to identify the minimum number of genetic alterations required to induce ALT. Notably, such other modifications are presumably either already present, or readily activated, in JFCF

and SW39 cells. Thus, I propose to generate conditionally-null *TERC*$^{F/F}$ cells in the HCT116 *ATRX*$^{Neo/o}$ background using CRISPR/Cas9 gene editing. Notably, *TERC*$^{-/-}$ cells cannot escape telomere crisis (this contrasts with the DN-hTERT technique I presented in **Chapter 5** where cells can readily escape crisis by telomerase over-expression or mutation of the integrated DN-hTERT construct) (Jones et al. 2014). HCT116 *ATRX*$^{Neo/o}$:*TERC*$^{F/F}$ cells would then be subjected to a CRISPR/Cas9 whole-genome gene knockout screen (Koike-Yusa *et al.* 2014). Notably, in this genetic screen, once the TERC gene is deleted the vast majority of cells will be negatively selected by the ensuing telomere crisis. Thus, any surviving clones would be of significant interest. These screens utilize a CRISPR/Cas9 lentivirus, in which the guide RNAs directing the Cas9 contain a barcode to report which gene they have inactivated (Koike-Yusa et al. 2014). Therefore, in any clone that does immortalize, the mutation responsible can easily be identified by PCR amplification of the lentiviral CRISPR and sequencing the resulting product. Mutation of the target gene of the CRISPR, and any likely off-target hits, could then be confirmed by PCR and sequencing. This approach would enable the identification of putative ALT-state regulatory genes. Candidate hits could then be validated by mutating these genes in a different telomerase-positive cancers, to ensure that the results weren't cell-line specific.

Sub-aim2: Discern the role of p53 in suppressing the ALT-specific chromatin mark, H3.3S31P.

Recent work has additionally described an ALT-specific genome histone mark, H3.3S31P (Chang et al. 2015), which is of significant interest as ATRX is a regulator of H3.3 genomic localization. Independently, the H3.3S31P modification is a global histone

modification that corresponds to chromosomal mis-segregation at mitosis, and triggers a p53-dependent cell cycle arrest (Hinchcliffe et al. 2016). The implication of these two findings is that p53 signaling may be somehow disrupted in ALT tumors, which has actually been commonly observed in ALT cells (Henson and Reddel 2010), but without true mechanistic insight. Since the H3.3S31P mark is rapidly induced during mitotic anaphase (Chang et al. 2015), the anaphase promotion complex (APC) may be the DNA damage sensor, as it is the primary mechanosensory complex that regulates chromosomal mis-segregation. Thus, my hypothetical model is that in ALT cells (which are frequently polypoid) the H3.3S31P histone mark is constantly activated. This signal is first recognized by the APC and transduced by some unknown factor to initiate p53-dependent senescence. Presumably, the easiest way for cells to escape this cellular checkpoint would be via p53 mutation, which perhaps explains the frequency of its mutation in ALT. In order to characterize this signal transduction network, I would propose to compare common $ALT^{p53\text{-}negative}$ cell lines to those rare ones which do express p53 ($ALT^{p53\text{-}postive}$) (Henson and Reddel 2010). Specifically, I would attempt to identify the H3.3S31P signal transduction pathway by comparing the proteins which interact with H3.3S31P in $ALT^{p53\text{-}postive}$ and $ALT^{p53\text{-}negative}$ cells. As a positive control, I would utilize primary, normal IMR90 fibroblasts (which presumably have a functional checkpoint) and initiate chromosome segregation using an established rapid chilling/rewarming protocol (Kasuboski *et al.* 2011). To identify this pathway in ALT cells, $ALT^{p53\text{-}negative}$ and $ALT^{p53\text{-}postive}$ cellular extracts would be collected, the protein:protein interactions would be crosslinked, followed by immunoprecipitation for H3.3S31P. Following a purification of the H3.3S31P-bound proteins, the crosslinking would be reversed. Differential binding partners would be detected by quantitative mass spectrometry-coupled proteomics. Any putative interactors

could be functionally confirmed by editing of the corresponding gene in an ALT$^{p53\text{-}negative}$ cell, followed by re-expression of p53. I would hypothesize that the absence of the H3.3S31-interacting proteins would render the ALT$^{p53\text{-}negative}$ cells not sensitive to the restoration of p53 expression. If successful, this would identify another regulatory network which is repressive of ALT that is likely independent of ATRX and thus would enhance the predictive power of using molecular pathology to identify the drivers of a patient's tumor. In the future, gene therapy of a patient's cancer designed to restore these mutated repressors of ALT could serve as an excellent treatment option.

3) How does ATRX repress ALT?

The work I present in **Chapter 4** and **Chapter 5** indicates that ATRX represses ALT initiation. At a molecular level, as *ATRX*-null HCT116 cells transit telomere crisis, there appears to be site-specific telomere-recombination events. While this suggests that *ATRX* is directly repressing recombination at telomeres, the molecular mechanism of this activity remains unknown. In addition to telomeric H3.3, ATRX also controls a portion of the H3.3 incorporation into various genomic regions in addition to telomeres. These regions are usually GC-rich and include: structural elements such as the pericentric chromatin (Filipescu *et al.* 2013), G-quadraplex-prone regions (Levy et al. 2015) and even endogenous retroelements (Goldberg et al. 2010). These series of genomic regions are intriguing as they represent areas that are all inherently recombinogenic, either due to the repetitive nature of the sequence, or for the likelihood of DNA damage to occur, as has been observed in telomeres (Sfeir et al. 2009). The objective of this Aim would be to determine if H3.3 serves an anti-recombinogenic purpose. To address this, I propose to first map the genomic distribution of ATRX-dependent H3.3, then determine if these

regions are resistant to gene conversion using gene editing technologies. These experiments would enable a mechanistic explanation for the apparent H3.3 enrichment in areas which are conceptually recombinogenic, but somehow repressed from such activity.

Sub-aim1: Map the Genomic Distribution of ATRX-dependent H3.3

To determine how genomic H3.3 influences recombination, the specific regions of the genome in which H3.3 is incorporated need to be identified. While this has been done previously in MEFs (Goldberg et al. 2010) and various mouse tissues (Levy et al. 2015), it has not been done in human cells. Unfortunately, given the difference between humans and mice with respect to telomere maintenance and the regulation of DNA repair, there is inherent risk in simply extrapolating the murine data to humans, and thus I propose that similar experiments need to be performed in human cells. Therefore, I would map the distribution of H3.3 throughout the genome by Chip-Seq in HCT116 and HCT116 $ATRX^{Neo/o}$ cells, as well as in the non-immortalized, untransformed, healthy human lung fibroblast lines, IMR90 and MRC5. In brief, this technique works by immunoprecipitation of H3.3 crosslinked to its cognate DNA, then the reversal of that crosslinking, and finally the identification of the H3.3-bound DNA sequence by next-generation sequencing. By performing this technique in these four different cell lines, we could determine where H3.3 can generally be found in the genome, what specific regions were ATRX-dependent, and if any of these regions were cell or cancer- specific.

Sub-aim2: Determine the effect of H3.3 on genomic gene conversion

Recent advances in gene editing technologies have rapidly expanded the ease of gene conversion studies in human cells. Notably, gene conversion between a genomic

site and a plasmid donor has been well established as a representative measurement of recombinogenic activity (Kan *et al.* 2014). Using the H3.3 distribution data from *SA3.1*, I propose to design multiple (~5) CRIPSR/Cas9 targeting constructs and plasmid donors which incorporate a promoter-driven *NEO* resistance gene into regions of the genome where H3.3 incorporation was ATRX-dependent in HCT116. 24 hours following the transfection of these targeting constructs into cells, I would single-cell subclone the transfected cells in drug-containing media. Subclones would be screened by PCR to identify those cells that were correctly targeted. The number of correctly targeted clones in each cell line would be representative of the relative amount of recombination potential at each locus. I hypothesize that the absence of ATRX-mediated H3.3 incorporation in these regions would facilitate an enhanced frequency of gene conversion at these sites. This would provide strong evidence that the genomic incorporation of H3.3 by ATRX serves to prevent recombination. Follow-up studies could address the mechanistic role of H3.3 in repressing recombination by examining potential differences in the recruitment of HDR-factors to these sites such as MRN, BRCA1, Rad51, *etc*., where I would hypothesize that CRISPR-mediated DNA breaks in H3.3-regions would be repressed for the active recruitment of such HDR-initiating factors. Such studies would begin to explain how ATRX-mediated H3.3 incorporation into the genome represses HDR at those specific sites.

Bibliography

Abascal, F., A. Corpet, Z. A. Gurard-Levin, D. Juan, F. Ochsenbein, D. Rico, A. Valencia and G. Almouzni (2013). "Subfunctionalization via adaptive evolution influenced by genomic context: the case of histone chaperones ASF1a and ASF1b." Mol Biol Evol **30**(8): 1853-1866.

Amorim, J. P., G. Santos, J. Vinagre and P. Soares (2016). "The role of ATRX in the alternative lengthening of telomeres (ALT) phenotype." Genes (Basel) **7**(9): 65-85.

Andres, S. N., A. Vergnes, D. Ristic, C. Wyman, M. Modesti and M. Junop (2012). "A human XRCC4-XLF complex bridges DNA." Nucleic Acids Res **40**(4): 1868-1878.

Arakawa, H., T. Bednar, M. Wang, K. Paul, E. Mladenov, A. A. Bencsik-Theilen and G. Iliakis (2012). "Functional redundancy between DNA ligases I and III in DNA replication in vertebrate cells." Nucleic Acids Res **40**(6): 2599-2610.

Audebert, M., B. Salles and P. Calsou (2004). "Involvement of poly(ADP-ribose) polymerase-1 and XRCC1/DNA ligase III in an alternative route for DNA double-strand breaks rejoining." J Biol Chem **279**(53): 55117-55126.

Audebert, M., B. Salles and P. Calsou (2008). "Effect of double-strand break DNA sequence on the PARP-1 NHEJ pathway." Biochem Biophys Res Commun **369**(3): 982-988.

Audebert, M., B. Salles, M. Weinfeld and P. Calsou (2006). "Involvement of polynucleotide kinase in a poly(ADP-ribose) polymerase-1-dependent DNA double-strand breaks rejoining pathway." J Mol Biol **356**(2): 257-265.

Azzalin, C. M., P. Reichenbach, L. Khoriauli, E. Giulotto and J. Lingner (2007). "Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends." Science **318**(5851): 798-801.

Badie, S., A. R. Carlos, C. Folio, K. Okamoto, P. Bouwman, J. Jonkers and M. Tarsounas (2015). "BRCA1 and CtIP promote alternative non-homologous end-joining at uncapped telomeres." EMBO J **34**(3): 410-424.

Baird, D. M., J. Rowson, D. Wynford-Thomas and D. Kipling (2003). "Extensive allelic variation and ultrashort telomeres in senescent human cells." Nat Genet **33**(2): 203-207.

Beneke, S., O. Cohausz, M. Malanga, P. Boukamp, F. Althaus and A. Burkle (2008). "Rapid regulation of telomere length is mediated by poly(ADP-ribose) polymerase-1." Nucleic Acids Res **36**(19): 6309-6317.

Bennardo, N., A. Cheng, N. Huang and J. M. Stark (2008). "Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair." PLoS Genet **4**(6): e1000110.

Benson, G. (1999). "Tandem repeats finder: a program to analyze DNA sequences." Nucleic Acids Res **27**(2): 573-580.

Bentley, D., J. Selfridge, J. K. Millar, K. Samuel, N. Hole, J. D. Ansell and D. W. Melton (1996). "DNA ligase I is required for fetal liver erythropoiesis but is not essential for mammalian cell viability." Nature Genetics **13**(4): 489-491.

Bentley, D. J., C. Harrison, A. M. Ketchen, N. J. Redhead, K. Samuel, M. Waterfall, J. D. Ansell and D. W. Melton (2002). "DNA ligase I null mouse cells show normal DNA repair activity but altered DNA replication and reduced genome stability." Journal of Cell Sci **115**(Pt 7): 1551-1561.

Bentley, J., C. P. Diggle, P. Harnden, M. A. Knowles and A. E. Kiltie (2004). "DNA double strand break repair in human bladder cancer is error prone and involves microhomology-associated end-joining." Nucleic Acids Res **32**(17): 5249-5259.

Betermier, M., P. Bertrand and B. S. Lopez (2014). "Is non-homologous end-joining really an inherently error-prone process?" Plos Genetics **10**(1): 1-9.

Blasco, M. A., H. W. Lee, M. P. Hande, E. Samper, P. M. Lansdorp, R. A. DePinho and C. W. Greider (1997). "Telomere shortening and tumor formation by mouse cells lacking telomerase RNA." Cell **91**(1): 25-34.

Bley, C. J., X. Qi, D. P. Rand, C. R. Borges, R. W. Nelson and J. J. Chen (2011). "RNA-protein binding interface in the telomerase ribonucleoprotein." Proc Natl Acad Sci U S A **108**(51): 20333-20338.

Blocher, D. and W. Pohlit (1982). "DNA double strand breaks in Ehrlich ascites tumour cells at low doses of x-rays. II. Can cell death be attributed to double strand breaks?" Int J Radiat Biol Relat Stud Phys Chem Med **42**(3): 329-338.

Boboila, C., F. W. Alt and B. Schwer (2012). "Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks." Advan in Immun **116**: 1-49.

Bock, F. J. and P. Chang (2016). "New directions in poly(ADP-ribose) polymerase biology." FEBS J **283**(22): 4017-4031.

Boltz, K. A., M. Jasti, J. M. Townley and D. E. Shippen (2014). "Analysis of poly(ADP-Ribose) polymerases in Arabidopsis telomere biology." PLoS One **9**(2): e88872.

Boulton, S. J. and S. P. Jackson (1996). "Saccharomyces cerevisiae Ku70 potentiates illegitimate DNA double-strand break repair and serves as a barrier to error-prone DNA repair pathways." EMBO J **15**(18): 5093-5103.

Boulton, S. J. and S. P. Jackson (1998). "Components of the Ku-dependent non-homologous end-joining pathway are involved in telomeric length maintenance and telomeric silencing." EMBO J **17**(6): 1819-1828.

Bower, K., C. E. Napier, S. L. Cole, R. A. Dagg, L. M. Lau, E. L. Duncan, E. L. Moy and R. R. Reddel (2012). "Loss of wild-type ATRX expression in somatic cell hybrids segregates with activation of Alternative Lengthening of Telomeres." PLoS One **7**(11): e50062.

Bradley, M. O. and K. W. Kohn (1979). "X-ray induced DNA double strand break production and repair in mammalian cells as measured by neutral filter elution." Nucleic Acids Res **7**(3): 793-804.

Bryan, T. M., A. Englezou, J. Gupta, S. Bacchetti and R. R. Reddel (1995). "Telomere elongation in immortal human cells without detectable telomerase activity." EMBO J **14**(17): 4240-4248.

Bryant, H. E., N. Schultz, H. D. Thomas, K. M. Parker, D. Flower, E. Lopez, S. Kyle, M. Meuth, N. J. Curtin and T. Helleday (2005). "Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase." Nature **434**(7035): 913-917.

Bunting, S. F., E. Callen, N. Wong, H. T. Chen, et al. (2010). "53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks." Cell **141**(2): 243-254.

Bunting, S. F. and A. Nussenzweig (2013). "End-joining, translocations and cancer." Nat Rev Canc **13**(7): 443-454.

Caldecott, K. W. (2007). "Mammalian single-strand break repair: mechanisms and links with chromatin." DNA Repair **6**(4): 443-453.
Caldecott, K. W. (2008). "Single-strand break repair and genetic disease." Nat Rev Genet **9**(8): 619-631.
Caldecott, K. W., C. K. McKeown, J. D. Tucker, S. Ljungquist and L. H. Thompson (1994). "An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III." Mol Cell Bio **14**(1): 68-76.
Cao, L., X. Xu, S. F. Bunting, J. Liu, et al. (2009). "A selective requirement for 53BP1 in the biological response to genomic instability induced by Brca1 deficiency." Mol Cell **35**(4): 534-541.
Capper, R., B. Britt-Compton, M. Tankimanova, J. Rowson, B. Letsolo, S. Man, M. Haughton and D. M. Baird (2007). "The nature of telomere fusion and a definition of the critical telomere length in human cells." Genes Dev **21**(19): 2495-2508.
Celli, G. B. and T. de Lange (2005). "DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion." Nat Cell Biol **7**(7): 712-718.
Celli, G. B., E. L. Denchi and T. de Lange (2006). "Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination." Nat Cell Biol **8**(8): 885-890.
Cesare, A. J. and R. R. Reddel (2010). "Alternative lengthening of telomeres: models, mechanisms and implications." Nat Rev Genet **11**(5): 319-330.
Chang, F. T., F. L. Chan, R. M. JD, M. Udugama, L. Mayne, P. Collas, J. R. Mann and L. H. Wong (2015). "CHK1-driven histone H3.3 serine 31 phosphorylation is important for chromatin maintenance and cell survival in human ALT cancer cells." Nucleic Acids Res **43**(5): 2603-2614.
Chang, H. H., G. Watanabe, C. A. Gerodimos, T. Ochi, T. L. Blundell, S. P. Jackson and M. R. Lieber (2016). "Different DNA end configurations dictate which NHEJ components are most important for joining efficiency." J Biol Chem **291**(47): 24377-24389.
Chapman, J. R., P. Barral, J. B. Vannier, V. Borel, M. Steger, A. Tomas-Loba, A. A. Sartori, I. R. Adams, F. D. Batista and S. J. Boulton (2013). "RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection." Mol Cell **49**(5): 858-871.
Chapman, J. R., A. J. Sossick, S. J. Boulton and S. P. Jackson (2012). "BRCA1-associated exclusion of 53BP1 from DNA damage sites underlies temporal control of DNA repair." J Cell Sci **125**(Pt 15): 3529-3534.
Chen, B., L. A. Gilbert, B. A. Cimini, J. Schnitzbauer, W. Zhang, G. W. Li, J. Park, E. H. Blackburn, J. S. Weissman, L. S. Qi and B. Huang (2013). "Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system." Cell **155**(7): 1479-1491.
Chen, J. L. and C. W. Greider (2003). "Determinants in mammalian telomerase RNA that mediate enzyme processivity and cross-species incompatibility." EMBO J **22**(2): 304-314.
Chen, X., A. Bahrami, A. Pappo, J. Easton, et al. (2014). "Recurrent somatic structural variations contribute to tumorigenesis in pediatric osteosarcoma." Cell Rep **7**(1): 104-112.

Chen, Y., Y. Yang, M. van Overbeek, J. R. Donigian, P. Baciu, T. de Lange and M. Lei (2008). "A shared docking motif in TRF1 and TRF2 used for differential recruitment of telomeric proteins." Science **319**(5866): 1092-1096.

Chen, Z., H. Yang and N. P. Pavletich (2008). "Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures." Nature **453**(7194): 489-484.

Cheng, Q., N. Barboule, P. Frit, D. Gomez, O. Bombarde, B. Couderc, G. S. Ren, B. Salles and P. Calsou (2011). "Ku counteracts mobilization of PARP1 and MRN in chromatin damaged with DNA double-strand breaks." Nucleic Acids Res **39**(22): 9605-9619.

Cheung, N. K., J. Zhang, C. Lu, M. Parker, et al. (2012). "Association of age at diagnosis and genetic mutations in patients with neuroblastoma." JAMA **307**(10): 1062-1071.

Ciccia, A. and S. J. Elledge (2010). "The DNA damage response: making it safe to play with knives." Mol Cell **40**(2): 179-204.

Cimprich, K. A. and D. Cortez (2008). "ATR: an essential regulator of genome integrity." Nat Rev Mol Cell Biol **9**(8): 616-627.

Clynes, D., C. Jelinska, B. Xella, H. Ayyub, C. Scott, M. Mitson, S. Taylor, D. R. Higgs and R. J. Gibbons (2015). "Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX." Nat Commun **6**: 7538: 1-11.

Colgin, L. M. and R. R. Reddel (1999). "Telomere maintenance mechanisms and cellular immortalization." Curr Opin Genet Dev **9**(1): 97-103.

Conomos, D., R. R. Reddel and H. A. Pickett (2014). "NuRD-ZNF827 recruitment to telomeres creates a molecular scaffold for homologous recombination." Nat Struct Mol Biol **21**(9): 760-770.

Costantini, S., L. Woodbine, L. Andreoli, P. A. Jeggo and A. Vindigni (2007). "Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK." DNA Repair (Amst) **6**(6): 712-722.

Cottarel, J., P. Frit, O. Bombarde, B. Salles, A. Negrel, S. Bernard, P. A. Jeggo, M. R. Lieber, M. Modesti and P. Calsou (2013). "A noncatalytic function of the ligation complex during nonhomologous end joining." J Cell Biol **200**(2): 173-186.

Counter, C. M., A. A. Avilion, C. E. LeFeuvre, N. G. Stewart, C. W. Greider, C. B. Harley and S. Bacchetti (1992). "Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity." EMBO J **11**(5): 1921-1929.

d'Adda di Fagagna, F., M. P. Hande, W. M. Tong, P. M. Lansdorp, Z. Q. Wang and S. P. Jackson (1999). "Functions of poly(ADP-ribose) polymerase in controlling telomere length and chromosomal stability." Nat Genet **23**(1): 76-80.

d'Adda di Fagagna, F., M. P. Hande, W. M. Tong, D. Roth, P. M. Lansdorp, Z. Q. Wang and S. P. Jackson (2001). "Effects of DNA nonhomologous end-joining factors on telomere length and chromosomal stability in mammalian cells." Curr Biol **11**(15): 1192-1196.

Dang, L. H., F. Chen, C. Ying, S. Y. Chun, S. A. Knock, H. D. Appelman and D. T. Dang (2006). "CDX2 has tumorigenic potential in the human colon cancer cell lines LOVO and SW48." Oncogene **25**(15): 2264-2272.

Dantzer, F., V. Schreiber, C. Niedergang, C. Trucco, E. Flatter, G. De La Rubia, J. Oliver, V. Rolli, J. Menissier-de Murcia and G. de Murcia (1999). "Involvement of poly(ADP-ribose) polymerase in base excision repair." Biochimie **81**(1-2): 69-75.

de Lange, T. (2010). "How shelterin solves the telomere end-protection problem." Cold Spring Harb Symp Quant Biol **75**: 167-177.
de Lange, T., L. Shiue, R. M. Myers, D. R. Cox, S. L. Naylor, A. M. Killery and H. E. Varmus (1990). "Structure and variability of human chromosome ends." Mol Cell Biol **10**(2): 518-527.
de Murcia, G. and J. Menissier de Murcia (1994). "Poly(ADP-ribose) polymerase: a molecular nick-sensor." Trends Biochem Sci **19**(4): 172-176.
de Murcia, J. M., C. Niedergang, C. Trucco, M. Ricoul, et al. (1997). "Requirement of poly(ADP-ribose) polymerase in recovery from DNA damage in mice and in cells." Proc Natl Acad Sci U S A **94**(14): 7303-7307.
de Sena-Tomas, C., E. Y. Yu, A. Calzada, W. K. Holloman, N. F. Lue and J. Perez-Martin (2015). "Fungal Ku prevents permanent cell cycle arrest by suppressing DNA damage signaling at telomeres." Nucleic Acids Res **43**(4): 2138-2151.
Decottignies, A. (2007). "Microhomology-mediated end joining in fission yeast is repressed by pku70 and relies on genes involved in homologous recombination." Genetics **176**(3): 1403-1415.
Decottignies, A. (2013). "Alternative end-joining mechanisms: a historical perspective." Front Genet **4**: 48.
Denchi, E. L. and T. de Lange (2007). "Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1." Nature **448**(7157): 1068-1071.
Deng, Z., J. Norseen, A. Wiedmer, H. Riethman and P. M. Lieberman (2009). "TERRA RNA binding to TRF2 facilitates heterochromatin formation and ORC recruitment at telomeres." Mol Cell **35**(4): 403-413.
DePristo, M. A., E. Banks, R. Poplin, K. V. Garimella, et al. (2011). "A framework for variation discovery and genotyping using next-generation DNA sequencing data." Nat Genet **43**(5): 491-498.
Dimitrova, N., Y. C. Chen, D. L. Spector and T. de Lange (2008). "53BP1 promotes non-homologous end joining of telomeres by increasing chromatin mobility." Nature **456**(7221): 524-528.
Doherty, J. E., L. E. Huye, K. Yusa, L. Zhou, N. L. Craig and M. H. Wilson (2012). "Hyperactive piggyBac gene transfer in human cells and in vivo." Hum Gene Ther **23**(3): 311-320.
Doherty, J. E., L. E. Huye, K. Yusa, L. Zhou, N. L. Craig and M. H. Wilson (2012). "Hyperactive piggyBac gene transfer in human cells and in vivo." Human Gene Therapy **23**(3): 311-320.
Doksani, Y. and T. de Lange (2014). "The role of double-strand break repair pathways at functional and dysfunctional telomeres." Cold Spring Harb Perspect Biol **6**(12): a016576.
Dunham, M. A., A. A. Neumann, C. L. Fasching and R. R. Reddel (2000). "Telomere maintenance by recombination in human cells." Nat Genet **26**(4): 447-450.
Eddy, B. E., G. S. Borman, G. E. Grubbs and R. D. Young (1962). "Identification of the oncogenic substance in rhesus monkey kidney cell culture as simian virus 40." Virology **17**: 65-75.
Edwards, S. L., R. Brough, C. J. Lord, R. Natrajan, R. Vatcheva, D. A. Levine, J. Boyd, J. S. Reis-Filho and A. Ashworth (2008). "Resistance to therapy caused by intragenic deletion in BRCA2." Nature **451**(7182): 1111-1115.

Eid, R., M. V. Demattei, H. Episkopou, C. Auge-Gouillou, A. Decottignies, N. Grandin and M. Charbonneau (2015). "genetic inactivation of ATRX leads to a decrease in the amount of telomeric cohesin and level of telomere transcription in human Glioma Cells." Mol Cell Biol **35**(16): 2818-2830.

el-Deiry, W. S., J. W. Harper, P. M. O'Connor, V. E. Velculescu, C. E. Canman, J. Jackman, J. A. Pietenpol, M. Burrell, D. E. Hill, Y. Wang and et al. (1994). "WAF1/CIP1 is induced in p53-mediated G1 arrest and apoptosis." Cancer Res **54**(5): 1169-1174.

Ellenberger, T. and A. E. Tomkinson (2008). "Eukaryotic DNA ligases: structural and functional insights." Annual Review of Biochemistry **77**: 313-338.

Epel, E. S., E. H. Blackburn, J. Lin, F. S. Dhabhar, N. E. Adler, J. D. Morrow and R. M. Cawthon (2004). "Accelerated telomere shortening in response to life stress." Proc Natl Acad Sci U S A **101**(49): 17312-17315.

Episkopou, H., I. Draskovic, A. Van Beneden, G. Tilman, M. Mattiussi, M. Gobin, N. Arnoult, A. Londono-Vallejo and A. Decottignies (2014). "Alternative Lengthening of Telomeres is characterized by reduced compaction of telomeric chromatin." Nucleic Acids Res **42**(7): 4391-4405.

Escribano-Diaz, C., A. Orthwein, A. Fradet-Turcotte, M. Xing, et al. (2013). "A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice." Mol Cell **49**(5): 872-883.

Farmer, H., N. McCabe, C. J. Lord, A. N. Tutt, et al. (2005). "Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy." Nature **434**(7035): 917-921.

Fattah, F., E. H. Lee, N. Weisensel, Y. Wang, N. Lichter and E. A. Hendrickson (2010). "Ku regulates the non-homologous end joining pathway choice of DNA double-strand break repair in human somatic cells." PLoS Genet **6**(2): e1000855.

Fattah, F. J., N. F. Lichter, K. R. Fattah, S. Oh and E. A. Hendrickson (2008). "Ku70, an essential gene, modulates the frequency of rAAV-mediated gene targeting in human somatic cells." Proc Natl Acad Sci U S A **105**(25): 8703-8708.

Feldser, D. M. and C. W. Greider (2007). "Short telomeres limit tumor progression in vivo by inducing senescence." Cancer Cell **11**(5): 461-469.

Filipescu, D., E. Szenker and G. Almouzni (2013). "Developmental roles of histone H3 variants and their chaperones." Trends Genet **29**(11): 630-640.

Fishbein, L., S. Khare, B. Wubbenhorst, D. DeSloover, et al. (2015). "Whole-exome sequencing identifies somatic ATRX mutations in pheochromocytomas and paragangliomas." Nat Commun **6**: 6140: 1-6.

Flynn, R. L., K. E. Cox, M. Jeitany, H. Wakimoto, et al. (2015). "Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors." Science **347**(6219): 273-277.

Fong, P. C., D. S. Boss, T. A. Yap, A. Tutt, et al. (2009). "Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers." N Engl J Med **361**(2): 123-134.

Frosina, G., P. Fortini, O. Rossi, F. Carrozzino, G. Raspaglio, L. S. Cox, D. P. Lane, A. Abbondandolo and E. Dogliotti (1996). "Two pathways for base excision repair in mammalian cells." J of Biol Chem **271**(16): 9573-9578.

Fujita, K., I. Horikawa, A. M. Mondal, L. M. Jenkins, E. Appella, B. Vojtesek, J. C. Bourdon, D. P. Lane and C. C. Harris (2010). "Positive feedback between p53 and TRF2

during telomere-damage signalling and cellular senescence." Nat Cell Biol **12**(12): 1205-1212.
Gao, Y., S. Katyal, Y. Lee, J. Zhao, J. E. Rehg, H. R. Russell and P. J. McKinnon (2011). "DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair." Nature **471**(7337): 240-244.
Garnett, M. J., E. J. Edelman, S. J. Heidorn, C. D. Greenman, et al. (2012). "Systematic identification of genomic markers of drug sensitivity in cancer cells." Nature **483**(7391): 570-575.
Gelmon, K. A., H. W. Hirte, A. Robidoux, K. S. Tonkin, M. Tischkowitz, K. Swenerton, D. Huntsman, J. Carmichael, E. Macpherson and A. M. Oza (2010). "Can we define tumors that will respond to PARP inhibitors? A phase II correlative study of olaparib in advanced serous ovarian cancer and triple-negative breast cancer." Clin Oncol **28**(15).
Ghezraoui, H., M. Piganeau, B. Renouf, J. B. Renaud, et al. (2014). "Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining." Mol Cell **55**(6): 829-842.
Ghosal, G. and J. Chen (2013). "DNA damage tolerance: a double-edged sword guarding the genome." Trans Canc Res **2**(3): 107-129.
Giannone, R. J., H. W. McDonald, G. B. Hurst, R. F. Shen, Y. Wang and Y. Liu (2010). "The protein network surrounding the human telomere repeat binding factors TRF1, TRF2, and POT1." PLoS One **5**(8): e12407.
Gibson, B. A. and W. L. Kraus (2012). "New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs." Nat Rev Mol Cell Biol **13**(7): 411-424.
Gilson, E. and V. Geli (2007). "How telomeres are replicated." Nat Rev Mol Cell Biol **8**(10): 825-838.
Girardi, A. J., D. Weinstein and P. S. Moorhead (1966). "SV40 transformation of human diploid cells. A parallel study of viral and karyologic parameters." Ann Med Exp Biol Fenn **44**(2): 242-254.
Godon, C., F. P. Cordelieres, D. Biard, N. Giocanti, F. Megnin-Chanet, J. Hall and V. Favaudon (2008). "PARP inhibition versus PARP-1 silencing: different outcomes in terms of single-strand break repair and radiation susceptibility." Nucleic Acids Res **36**(13): 4454-4464.
Goldberg, A. D., L. A. Banaszynski, K. M. Noh, P. W. Lewis, et al. (2010). "Distinct factors control histone variant H3.3 localization at specific genomic regions." Cell **140**(5): 678-691.
Goldstein, S. (1990). "Replicative senescence: the human fibroblast comes of age." Science **249**(4973): 1129-1133.
Gomez, M., J. Wu, V. Schreiber, J. Dunlap, F. Dantzer, Y. Wang and Y. Liu (2006). "PARP1 Is a TRF2-associated poly(ADP-ribose)polymerase and protects eroded telomeres." Mol Biol Cell **17**(4): 1686-1696.
Gottlich, B., S. Reichenberger, E. Feldmann and P. Pfeiffer (1998). "Rejoining of DNA double-strand breaks in vitro by single-strand annealing." Eur J Bioch/ FEBS **258**(2): 387-395.
Greider, C. W. (1991). "Telomerase is processive." Mol Cell Biol **11**(9): 4572-4580.
Greider, C. W. and E. H. Blackburn (1985). "Identification of a specific telomere terminal transferase activity in Tetrahymena extracts." Cell **43**(2 Pt 1): 405-413.

Gu, H., J. D. Marth, P. C. Orban, H. Mossmann and K. Rajewsky (1994). "Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting." Science **265**(5168): 103-106.
Guirouilh-Barbat, J., E. Rass, I. Plo, P. Bertrand and B. S. Lopez (2007). "Defects in XRCC4 and KU80 differentially affect the joining of distal nonhomologous ends." Proc Natl Acad of Sci **104**(52): 20902-20907.
Guo, Y., M. Kartawinata, J. Li, H. A. Pickett, et al. (2014). "Inherited bone marrow failure associated with germline mutation of ACD, the gene encoding telomere protein TPP1." Blood **124**(18): 2767-2774.
Haber, J. E., P. C. Thorburn and D. Rogers (1984). "Meiotic and mitotic behavior of dicentric chromosomes in Saccharomyces cerevisiae." Genetics **106**(2): 185-205.
Hahn, W. C., S. A. Stewart, M. W. Brooks, S. G. York, E. Eaton, A. Kurachi, R. L. Beijersbergen, J. H. Knoll, M. Meyerson and R. A. Weinberg (1999). "Inhibition of telomerase limits the growth of human cancer cells." Nat Med **5**(10): 1164-1170.
Hanahan, D. and R. A. Weinberg (2000). "The hallmarks of cancer." Cell **100**(1): 57-70.
Hanahan, D. and R. A. Weinberg (2011). "Hallmarks of cancer: the next generation." Cell **144**(5): 646-674.
Haradhvala, N. J., P. Polak, P. Stojanov, K. R. Covington, et al. (2016). "Mutational strand asymmetries in cancer genomes reveal mechanisms of DNA damage and repair." Cell **164**(3): 538-549.
Harley, C. B., A. B. Futcher and C. W. Greider (1990). "Telomeres shorten during ageing of human fibroblasts." Nature **345**(6274): 458-460.
Hartlerode, A. J. and R. Scully (2009). "Mechanisms of double-strand break repair in somatic mammalian cells." The Biochem J **423**(2): 157-168.
Hassa, P. O. and M. O. Hottiger (2008). "The diverse biological roles of mammalian PARPS, a small but powerful family of poly-ADP-ribose polymerases." Front Biosci **13**: 3046-3082.
Hayaishi, O. and K. Ueda (1977). "Poly(ADP-ribose) and ADP-ribosylation of proteins." Annu Rev Biochem **46**: 95-116.
Hayflick, L. and P. S. Moorhead (1961). "The serial cultivation of human diploid cell strains." Exp Cell Res **25**: 585-621.
Heaphy, C. M., R. F. de Wilde, Y. Jiao, A. P. Klein, et al. (2011). "Altered telomeres in tumors with ATRX and DAXX mutations." Science **333**(6041): 425.
Heaphy, C. M., A. P. Subhawong, S. M. Hong, M. G. Goggins, et al. (2011). "Prevalence of the alternative lengthening of telomeres telomere maintenance mechanism in human cancer subtypes." Am J Pathol **179**(4): 1608-1615.
Helleday, T. (2011). "The underlying mechanism for the PARP and BRCA synthetic lethality: clearing up the misunderstandings." Mol Oncol **5**(4): 387-393.
Helleday, T., S. Eshtad and S. Nik-Zainal (2014). "Mechanisms underlying mutational signatures in human cancers." Nat Rev Genet **15**(9): 585-598.
Hendrickson, E. A. (2008). Gene targeting in human somatic cells. Source Book of Models for Biomedical Research. P. M. Conn. Totowa, NJ, Humana Press, Inc.**:** 509-525.
Hendrickson, E. A. and D. M. Baird (2015). "Alternative end joining, clonal evolution, and escape from a telomere-driven crisis." Mol Cell Oncol **2**(1): e975623.
Henson, J. D., Y. Cao, L. I. Huschtscha, A. C. Chang, A. Y. Au, H. A. Pickett and R. R. Reddel (2009). "DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity." Nat Biotechnol **27**(12): 1181-1185.

Henson, J. D., A. A. Neumann, T. R. Yeager and R. R. Reddel (2002). "Alternative lengthening of telomeres in mammalian cells." Oncogene **21**(4): 598-610.
Henson, J. D. and R. R. Reddel (2010). "Assaying and investigating Alternative Lengthening of Telomeres activity in human cells and cancers." FEBS Lett **584**(17): 3800-3811.
Hinchcliffe, E. H., C. A. Day, K. B. Karanjeet, S. Fadness, A. Langfald, K. T. Vaughan and Z. Dong (2016). "Chromosome missegregation during anaphase triggers p53 cell cycle arrest through histone H3.3 Ser31 phosphorylation." Nat Cell Biol **18**(6): 668-675.
Hirata, R., J. Chamberlain, R. Dong and D. W. Russell (2002). "Targeted transgene insertion into human chromosomes by adeno-associated virus vectors." Nat Biotechnol **20**(7): 735-738.
Hladnik, U., W. L. Nyhan and M. Bertelli (2008). "Variable expression of HPRT deficiency in 5 members of a family with the same mutation." Archiv Neur **65**(9): 1240-1243.
Hochegger, H., D. Dejsuphong, T. Fukushima, C. Morrison, et al. (2006). "Parp-1 protects homologous recombination from interference by Ku and Ligase IV in vertebrate cells." EMBO J **25**(6): 1305-1314.
Houchmandzadeh, B., J. F. Marko, D. Chatenay and A. Libchaber (1997). "Elasticity and structure of eukaryote chromosomes studied by micromanipulation and micropipette aspiration." J Cell Biol **139**(1): 1-12.
Hu, Y., G. Shi, L. Zhang, F. Li, Y. Jiang, S. Jiang, W. Ma, Y. Zhao, Z. Songyang and J. Huang (2016). "Switch telomerase to ALT mechanism by inducing telomeric DNA damages and dysfunction of ATRX and DAXX." Sci Rep **6**: 32280.
Huang, Y. and L. Li (2013). "DNA crosslinking damage and cancer - a tale of friend and foe." Transl Canc Res **2**(3): 144-154.
Huschtscha, L. I. and R. Holliday (1983). "Limited and unlimited growth of SV40-transformed cells from human diploid MRC-5 fibroblasts." J Cell Sci **63**: 77-99.
Huyen, Y., O. Zgheib, R. A. Ditullio, Jr., V. G. Gorgoulis, P. Zacharatos, T. J. Petty, E. A. Sheston, H. S. Mellert, E. S. Stavridi and T. D. Halazonetis (2004). "Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks." Nature **432**(7015): 406-411.
Hwang, W. Y., Y. Fu, D. Reyon, M. L. Maeder, S. Q. Tsai, J. D. Sander, R. T. Peterson, J. R. Yeh and J. K. Joung (2013). "Efficient genome editing in zebrafish using a CRISPR-Cas system." Nat Biotechnol **31**(3): 227-229.
Ide, T., Y. Tsuji, T. Nakashima and S. Ishibashi (1984). "Progress of aging in human diploid cells transformed with a tsA mutant of simian virus 40." Exp Cell Res **150**(2): 321-328.
Iliakis, G. (2009). "Backup pathways of NHEJ in cells of higher eukaryotes: cell cycle dependence." Radiother Oncol **92**(3): 310-315.
Ira, G., D. Satory and J. E. Haber (2006). "Conservative inheritance of newly synthesized DNA in double-strand break-induced gene conversion." Mol Cell Biol **26**(24): 9424-9429.
Ivanauskiene, K., E. Delbarre, J. D. McGhie, T. Kuntziger, L. H. Wong and P. Collas (2014). "The PML-associated protein DEK regulates the balance of H3.3 loading on chromatin and is important for telomere integrity." Genome Res **24**(10): 1584-1594.

Janson, C., K. Nyhan and J. P. Murnane (2015). "Replication stress and telomere dysfunction are present in cultured human embryonic stem cells." Cytogenet Genome Res **146**(4): 251-260.
Jasin, M. (1996). "Genetic manipulation of genomes with rare-cutting endonucleases." Trends Genet **12**(6): 224-228.
Jasin, M. and R. Rothstein (2013). "Repair of strand breaks by homologous recombination." Cold Spring Harb Perspect Biol **5**(11): a012740.
Jaspers, J. E., A. Kersbergen, U. Boon, W. Sol, et al. (2013). "Loss of 53BP1 causes PARP inhibitor resistance in Brca1-mutated mouse mammary tumors." Cancer Discov **3**(1): 68-81.
Jee, J., A. Rasouly, I. Shamovsky, Y. Akivis, S. R. Steinman, B. Mishra and E. Nudler (2016). "Rates and mechanisms of bacterial mutagenesis from maximum-depth sequencing." Nature **534**(7609): 693-696.
Jiang, W. Q., A. Nguyen, Y. Cao, A. C. Chang and R. R. Reddel (2011). "HP1-mediated formation of alternative lengthening of telomeres-associated PML bodies requires HIRA but not ASF1a." PLoS One **6**(2): e17036.
Jiang, W. Q., Z. H. Zhong, J. D. Henson, A. A. Neumann, A. C. Chang and R. R. Reddel (2005). "Suppression of alternative lengthening of telomeres by Sp100-mediated sequestration of the MRE11/RAD50/NBS1 complex." Mol Cell Biol **25**(7): 2708-2721.
Jiao, Y., C. Shi, B. H. Edil, R. F. de Wilde, et al. (2011). "DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors." Science **331**(6021): 1199-1203.
Jones, R. E., S. Oh, J. W. Grimstead, J. Zimbric, L. Roger, N. H. Heppel, K. E. Ashelford, K. Liddiard, E. A. Hendrickson and D. M. Baird (2014). "Escape from telomere-driven crisis is DNA ligase III dependent." Cell Rep **8**(4): 1063-1076.
Ju, Z. and K. Lenhard Rudolph (2008). "Telomere dysfunction and stem cell ageing." Biochimie **90**(1): 24-32.
Kabotyanski, E. B., L. Gomelsky, J. O. Han, T. D. Stamato and D. B. Roth (1998). "Double-strand break repair in Ku86- and XRCC4-deficient cells." Nucleic Acids Res **26**(23): 5333-5342.
Kan, Y., B. Ruis, S. Lin and E. A. Hendrickson (2014). "The mechanism of gene targeting in human somatic cells." PLoS Genet **10**(4): e1004251.
Karlseder, J., D. Broccoli, Y. Dai, S. Hardy and T. de Lange (1999). "p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2." Science **283**(5406): 1321-1325.
Kass, E. M. and M. Jasin (2010). "Collaboration and competition between DNA double-strand break repair pathways." FEBS letters **584**(17): 3703-3708.
Kasuboski, J. M., J. R. Bader, P. S. Vaughan, S. B. Tauhata, et al. (2011). "Zwint-1 is a novel Aurora B substrate required for the assembly of a dynein-binding platform on kinetochores." Mol Biol Cell **22**(18): 3318-3330.
Katyal, S. and P. J. McKinnon (2011). "Disconnecting XRCC1 and DNA ligase III." Cell cycle **10**(14): 2269-2275.
Kaul, Z., A. J. Cesare, L. I. Huschtscha, A. A. Neumann and R. R. Reddel (2011). "Five dysfunctional telomeres predict onset of senescence in human cells." EMBO Rep **13**(1): 52-59.

Keegan, C. E., J. E. Hutz, T. Else, M. Adamska, S. P. Shah, A. E. Kent, J. M. Howes, W. G. Beamer and G. D. Hammer (2005). "Urogenital and caudal dysgenesis in adrenocortical dysplasia (acd) mice is caused by a splicing mutation in a novel telomeric regulator." Hum Mol Genet **14**(1): 113-123.
Khan, I. F., R. K. Hirata and D. W. Russell (2011). "AAV-mediated gene targeting methods for human cells." Nat Protoc **6**(4): 482-501.
Kleine, H. and B. Luscher (2009). "Learning how to read ADP-ribosylation." Cell **139**(1): 17-19.
Knight, R., P. Maxwell, A. Birmingham, J. Carnes, et al. (2007). "PyCogent: a toolkit for making sense from sequence." Genome Biol **8**(8): R171.
Kohli, M., C. Rago, C. Lengauer, K. W. Kinzler and B. Vogelstein (2004). "Facile methods for generating human somatic cell gene knockouts using recombinant adeno-associated viruses." Nucleic Acids Res **32**(1): e3.
Koike-Yusa, H., Y. Li, E. P. Tan, C. Velasco-Herrera Mdel and K. Yusa (2014). "Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library." Nat Biotechnol **32**(3): 267-273.
Koschmann, C., P. R. Lowenstein and M. G. Castro (2016). "ATRX mutations and glioblastoma: Impaired DNA damage repair, alternative lengthening of telomeres, and genetic instability." Mol Cell Oncol **3**(3): e1167158.
Krishnakumar, R. and W. L. Kraus (2010). "The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets." Mol Cell **39**(1): 8-24.
Krokan, H. E. and M. Bjoras (2013). "Base excision repair." Cold Spring Harbor Persp Biol **5**(4): a012583.
Kuzminov, A. (2001). "Single-strand interruptions in replicating chromosomes cause double-strand breaks." Proc Natl Acad Sci U S A **98**(15): 8241-8246.
Lafferty-Whyte, K., C. J. Cairney, M. B. Will, N. Serakinci, M. G. Daidone, N. Zaffaroni, A. Bilsland and W. N. Keith (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): 3765-3774.
Lakshmipathy, U. and C. Campbell (1999). "The human DNA ligase III gene encodes nuclear and mitochondrial proteins." Mol Cell Biol **19**(5): 3869-3876.
Langerak, P., E. Mejia-Ramirez, O. Limbo and P. Russell (2011). "Release of Ku and MRN from DNA ends by Mre11 nuclease activity and Ctp1 is required for homologous recombination repair of double-strand breaks." PLoS Genet **7**(9): e1002271.
Law, M. J., K. M. Lower, H. P. Voon, J. R. Hughes, et al. (2010). "ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner." Cell **143**(3): 367-378.
Le Chalony, C., F. Hoffschir, L. R. Gauthier, J. Gross, D. S. Biard, F. D. Boussin and V. Pennaneach (2012). "Partial complementation of a DNA ligase I deficiency by DNA ligase III and its impact on cell survival and telomere stability in mammalian cells." Cell Mol Life Sci : CMLS **69**(17): 2933-2949.
Leahy, J. J., B. T. Golding, R. J. Griffin, I. R. Hardcastle, C. Richardson, L. Rigoreau and G. C. Smith (2004). "Identification of a highly potent and selective DNA-dependent protein kinase (DNA-PK) inhibitor (NU7441) by screening of chromenone libraries." Bioorg Med Chem Lett **14**(24): 6083-6087.
Lebofsky, R. and A. Bensimon (2003). "Single DNA molecule analysis: applications of molecular combing." Brief Funct Genomic Proteomic **1**(4): 385-396.

Lee, O. H., H. Kim, Q. He, H. J. Baek, D. Yang, L. Y. Chen, J. Liang, H. K. Chae, A. Safari, D. Liu and Z. Songyang (2011). "Genome-wide YFP fluorescence complementation screen identifies new regulators for telomere signaling in human cells." Mol Cell Proteomics **10**(2): M110 001628.

Lee, S. S., C. Bohrson, A. M. Pike, S. J. Wheelan and C. W. Greider (2015). "ATM kinase is required for telomere elongation in mouse and human cells." Cell Rep **13**(8): 1623-1632.

Lescale, C., H. Lenden Hasse, A. N. Blackford, G. Balmus, et al. (2016). "Specific Roles of XRCC4 Paralogs PAXX and XLF during V(D)J Recombination." Cell Rep **16**(11): 2967-2979.

Leung, J. W., G. Ghosal, W. Wang, X. Shen, J. Wang, L. Li and J. Chen (2013). "Alpha thalassemia/mental retardation syndrome X-linked gene product ATRX is required for proper replication restart and cellular resistance to replication stress." J Biol Chem **288**(9): 6342-6350.

Levy, M. A., K. D. Kernohan, Y. Jiang and N. G. Berube (2015). "ATRX promotes gene expression by facilitating transcriptional elongation through guanine-rich coding regions." Hum Mol Genet **24**(7): 1824-1835.

Lewis, P. W., S. J. Elsaesser, K. M. Noh, S. C. Stadler and C. D. Allis (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): 14075-14080.

Li, G., C. Nelsen and E. A. Hendrickson (2002). "Ku86 is essential in human somatic cells." Proc Natl Acad Sci U S A **99**(2): 832-837.

Li, H. and R. Durbin (2009). "Fast and accurate short read alignment with Burrows-Wheeler transform." Bioinformatics **25**(14): 1754-1760.

Li, Y., C. Schwab, S. L. Ryan, E. Papaemmanuil, et al. (2014). "Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia." Nature **508**(7494): 98-102.

Liang, L., L. Deng, S. C. Nguyen, X. Zhao, C. D. Maulion, C. Shao and J. A. Tischfield (2008). "Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks." Nucleic Acids Res **36**(10): 3297-3310.

Liddiard, K., B. Ruis, T. Takasugi, A. Harvey, K. E. Ashelford, E. A. Hendrickson and D. M. Baird (2016). "Sister chromatid telomere fusions, but not NHEJ-mediated inter-chromosomal telomere fusions, occur independently of DNA ligases 3 and 4." Genome Res **26**(5): 588-600.

Lieber, M. R. (2010). "The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway." Annu Rev Biochem **79**: 181-211.

Lieber, M. R., Y. Ma, U. Pannicke and K. Schwarz (2003). "Mechanism and regulation of human non-homologous DNA end-joining." Nat Rev Mol Cell Biol **4**(9): 712-720.

Lin, T. T., B. T. Letsolo, R. E. Jones, J. Rowson, G. Pratt, S. Hewamana, C. Fegan, C. Pepper and D. M. Baird (2010). "Telomere dysfunction and fusion during the progression of chronic lymphocytic leukemia: evidence for a telomere crisis." Blood **116**(11): 1899-1907.

Lio, P. (2003). "Wavelets in bioinformatics and computational biology: state of art and perspectives." Bioinformatics **19**(1): 2-9.

Liu, D., A. Safari, M. S. O'Connor, D. W. Chan, A. Laegeler, J. Qin and Z. Songyang (2004). "PTOP interacts with POT1 and regulates its localization to telomeres." Nat Cell Biol **6**(7): 673-680.

Lopez, V., N. Barinova, M. Onishi, S. Pobiega, J. R. Pringle, K. Dubrana and S. Marcand (2015). "Cytokinesis breaks dicentric chromosomes preferentially at pericentromeric regions and telomere fusions." Genes Dev **29**(3): 322-336.

Lord, C. J. and A. Ashworth (2013). "Mechanisms of resistance to therapies targeting BRCA-mutant cancers." Nat Med **19**(11): 1381-1388.

Lovejoy, C. A., W. Li, S. Reisenweber, S. Thongthip, 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): e1002772.

Ludwig, T., D. L. Chapman, V. E. Papaioannou and A. Efstratiadis (1997). "Targeted mutations of breast cancer susceptibility gene homologs in mice: lethal phenotypes of Brca1, Brca2, Brca1/Brca2, Brca1/p53, and Brca2/p53 nullizygous embryos." Genes Dev **11**(10): 1226-1241.

Lue, N. F. (2005). "A physical and functional constituent of telomerase anchor site." J Biol Chem **280**(28): 26586-26591.

Lue, N. F. and Z. Li (2007). "Modeling and structure function analysis of the putative anchor site of yeast telomerase." Nucleic Acids Res **35**(15): 5213-5222.

Lukashchuk, V. and R. D. Everett (2010). "Regulation of ICP0-null mutant herpes simplex virus type 1 infection by ND10 components ATRX and hDaxx." J Virol **84**(8): 4026-4040.

Lundblad, V. and J. W. Szostak (1989). "A mutant with a defect in telomere elongation leads to senescence in yeast." Cell **57**(4): 633-643.

Luo, Y., L. Lin, L. Bolund and C. B. Sorensen (2014). "Efficient construction of rAAV-based gene targeting vectors by Golden Gate cloning." Biotechniques **56**(5): 263-268.

Maciejowski, J., Y. Li, N. Bosco, P. J. Campbell and T. de Lange (2015). "Chromothripsis and Kataegis Induced by Telomere Crisis." Cell **163**(7): 1641-1654.

Mahaney, B. L., M. Hammel, K. Meek, J. A. Tainer and S. P. Lees-Miller (2013). "XRCC4 and XLF form long helical protein filaments suitable for DNA end protection and alignment to facilitate DNA double strand break repair." Bioch Cell Biol **91**(1): 31-41.

Mangerich, A. and A. Burkle (2012). "Pleiotropic cellular functions of PARP1 in longevity and aging: genome maintenance meets inflammation." Oxid Med Cell Longev **2012**: 321653.

Mansour, W. Y., S. Schumacher, R. Rosskopf, T. Rhein, F. Schmidt-Petersen, F. Gatzemeier, F. Haag, K. Borgmann, H. Willers and J. Dahm-Daphi (2008). "Hierarchy of nonhomologous end-joining, single-strand annealing and gene conversion at site-directed DNA double-strand breaks." Nucleic Acids Res **36**(12): 4088-4098.

Mao, Z., M. Bozzella, A. Seluanov and V. Gorbunova (2008). "DNA repair by nonhomologous end joining and homologous recombination during cell cycle in human cells." Cell cycle **7**(18): 2902-2906.

Martin, G. M. (1978). "Genetic syndromes in man with potential relevance to the pathobiology of aging." Birth Defects Orig Artic Ser **14**(1): 5-39.

Martin, G. M., A. C. Smith, D. J. Ketterer, C. E. Ogburn and C. M. Disteche (1985). "Increased chromosomal aberrations in first metaphases of cells isolated from the kidneys of aged mice." Isr J Med Sci **21**(3): 296-301.

Martinez, P. and M. A. Blasco (2010). "Role of shelterin in cancer and aging." Aging Cell **9**(5): 653-666.

Martinez, P. and M. A. Blasco (2011). "Telomeric and extra-telomeric roles for telomerase and the telomere-binding proteins." Nat Rev Cancer **11**(3): 161-176.

Marzec, P., C. Armenise, G. Perot, F. M. Roumelioti, E. Basyuk, S. Gagos, F. Chibon and J. Dejardin (2015). "Nuclear-receptor-mediated telomere insertion leads to genome instability in ALT cancers." Cell **160**(5): 913-927.

Mattarucchi, E., V. Guerini, A. Rambaldi, L. Campiotti, A. Venco, F. Pasquali, F. Lo Curto and G. Porta (2008). "Microhomologies and interspersed repeat elements at genomic breakpoints in chronic myeloid leukemia." Genes Chromo Cancer **47**(7): 625-632.

McClintock, B. (1939). "The behavior in successive nuclear divisions of a chromosome broken at meiosis." Proc Natl Acad Sci U S A **25**(8): 405-416.

McClintock, B. (1941). "The Stability of Broken Ends of Chromosomes in Zea Mays." Genetics **26**(2): 234-282.

McKenna, A., M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly and M. A. DePristo (2010). "The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data." Genome Res **20**(9): 1297-1303.

McVey, M. and S. E. Lee (2008). "MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings." Trends in genetics : TIG **24**(11): 529-538.

Michishita, E., R. A. McCord, E. Berber, M. Kioi, et al. (2008). "SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin." Nature **452**(7186): 492-496.

Miller, D. G., G. D. Trobridge, L. M. Petek, M. A. Jacobs, R. Kaul and D. W. Russell (2005). "Large-scale analysis of adeno-associated virus vector integration sites in normal human cells." J Virol **79**(17): 11434-11442.

Miller, K. M., J. V. Tjeertes, J. Coates, G. Legube, S. E. Polo, S. Britton and S. P. Jackson (2010). "Human HDAC1 and HDAC2 function in the DNA-damage response to promote DNA nonhomologous end-joining." Nat Struct Mol Biol **17**(9): 1144-1151.

Mimori, T. and J. A. Hardin (1986). "Mechanism of interaction between Ku protein and DNA." J Biol Chem **261**(22): 10375-10379.

Min, J., W. E. Wright and J. W. Shay (2017). "Alternative lengthening of telomeres can be maintained by preferential elongation of lagging strands." Nucleic Acids Res. **45** (5): 2615-2628.

Min, W., C. Bruhn, P. Grigaravicius, Z. W. Zhou, et al. (2013). "Poly(ADP-ribose) binding to Chk1 at stalled replication forks is required for S-phase checkpoint activation." Nat Commun **4**: 2993-3007.

Miwa, M., S. Hanai, P. Poltronieri, M. Uchida and K. Uchida (1999). "Functional analysis of poly(ADP-ribose) polymerase in Drosophila melanogaster." Mol Cell Biochem **193**(1-2): 103-107.

Mladenov, E. and G. Iliakis (2011). "Induction and repair of DNA double strand breaks: the increasing spectrum of non-homologous end joining pathways." Mutation Res **711**(1-2): 61-72.
Morin, G. B. (1989). "The human telomere terminal transferase enzyme is a ribonucleoprotein that synthesizes TTAGGG repeats." Cell **59**(3): 521-529.
Moynahan, M. E. and M. Jasin (1997). "Loss of heterozygosity induced by a chromosomal double-strand break." Proc Natl Acad Sci U S A **94**(17): 8988-8993.
Murai, J., S. Y. Huang, B. B. Das, A. Renaud, Y. Zhang, J. H. Doroshow, J. Ji, S. Takeda and Y. Pommier (2012). "Trapping of PARP1 and PARP2 by clinical PARP inhibitors." Cancer Res **72**(21): 5588-5599.
Muramatsu, M., K. Kinoshita, S. Fagarasan, S. Yamada, Y. Shinkai and T. Honjo (2000). "Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme." Cell **102**(5): 553-563.
Murnane, J. P. (2010). "Telomere loss as a mechanism for chromosome instability in human cancer." Cancer Res **70**(11): 4255-4259.
Nakai, H., X. Wu, S. Fuess, T. A. Storm, D. Munroe, E. Montini, S. M. Burgess, M. Grompe and M. A. Kay (2005). "Large-scale molecular characterization of adeno-associated virus vector integration in mouse liver." J Virol **79**(6): 3606-3614.
Nandakumar, J., C. F. Bell, I. Weidenfeld, A. J. Zaug, L. A. Leinwand and T. R. Cech (2012). "The TEL patch of telomere protein TPP1 mediates telomerase recruitment and processivity." Nature **492**(7428): 285-289.
Nandakumar, J. and T. R. Cech (2013). "Finding the end: recruitment of telomerase to telomeres." Nat Rev Mol Cell Biol **14**(2): 69-82.
Napier, C. E., L. I. Huschtscha, A. Harvey, K. Bower, J. R. Noble, E. A. Hendrickson and R. R. Reddel (2015). "ATRX represses alternative lengthening of telomeres." Oncotarget **6**(18): 16543-16558.
Negrini, S., V. G. Gorgoulis and T. D. Halazonetis (2010). "Genomic instability--an evolving hallmark of cancer." Nat Rev Mol Cell Biol **11**(3): 220-228.
Nussenzweig, A. and M. C. Nussenzweig (2007). "A backup DNA repair pathway moves to the forefront." Cell **131**(2): 223-225.
O'Sullivan, R. J. and G. Almouzni (2014). "Assembly of telomeric chromatin to create ALTernative endings." Trends Cell Biol **24**(11): 675-685.
O'Sullivan, R. J., N. Arnoult, D. H. Lackner, L. Oganesian, C. Haggblom, A. Corpet, G. Almouzni and J. Karlseder (2014). "Rapid induction of alternative lengthening of telomeres by depletion of the histone chaperone ASF1." Nat Struct Mol Biol **21**(2): 167-174.
O'Sullivan, R. J. and J. Karlseder (2010). "Telomeres: protecting chromosomes against genome instability." Nat Rev Mol Cell Biol **11**(3): 171-181.
Ochi, T., A. N. Blackford, J. Coates, S. Jhujh, et al. (2015). "DNA repair. PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair." Science **347**(6218): 185-188.
Oh, S., A. Harvey, J. Zimbric, Y. Wang, T. Nguyen, P. J. Jackson and E. A. Hendrickson (2014). "DNA ligase III and DNA ligase IV carry out genetically distinct forms of end joining in human somatic cells." DNA Repair (Amst) **21**: 97-110.
Oh, S., Y. Wang, J. Zimbric and E. A. Hendrickson (2013). "Human LIGIV is synthetically lethal with the loss of Rad54B-dependent recombination and is required for certain

chromosome fusion events induced by telomere dysfunction." Nucleic Acids Res **41**(3): 1734-1749.
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): 181-190.
Pacchierotti, F. and R. Ranaldi (2006). "Mechanisms and risk of chemically induced aneuploidy in mammalian germ cells." Curr Pharm Des **12**(12): 1489-1504.
Palm, W. and T. de Lange (2008). "How shelterin protects mammalian telomeres." Annu Rev Genet **42**: 301-334.
Parkinson, G. N., M. P. Lee and S. Neidle (2002). "Crystal structure of parallel quadruplexes from human telomeric DNA." Nature **417**(6891): 876-880.
Paul, K., M. Wang, E. Mladenov, A. Bencsik-Theilen, T. Bednar, W. Wu, H. Arakawa and G. Iliakis (2013). "DNA ligases I and III cooperate in alternative non-homologous end-joining in vertebrates." PLoS One **8**(3): e59505.
Pereira-Smith, O. M. and J. R. Smith (1988). "Genetic analysis of indefinite division in human cells: identification of four complementation groups." Proc Natl Acad Sci U S A **85**(16): 6042-6046.
Perez-Jannotti, R. M., S. M. Klein and D. F. Bogenhagen (2001). "Two forms of mitochondrial DNA ligase III are produced in Xenopus laevis oocytes." J Biol Chem **276**(52): 48978-48987.
Perrem, K., T. M. Bryan, A. Englezou, T. Hackl, E. L. Moy and R. R. Reddel (1999). "Repression of an alternative mechanism for lengthening of telomeres in somatic cell hybrids." Oncogene **18**(22): 3383-3390.
Pickett, H. A. and R. R. Reddel (2015). "Molecular mechanisms of activity and derepression of alternative lengthening of telomeres." Nat Struct Mol Biol **22**(11): 875-880.
Poulet, A., R. Buisson, C. Faivre-Moskalenko, M. Koelblen, et al. (2009). "TRF2 promotes, remodels and protects telomeric Holliday junctions." EMBO J **28**(6): 641-651.
Preto, A., S. K. Singhrao, M. F. Haughton, D. Kipling, D. Wynford-Thomas and C. J. Jones (2004). "Telomere erosion triggers growth arrest but not cell death in human cancer cells retaining wild-type p53: implications for antitelomerase therapy." Oncogene **23**(23): 4136-4145.
Prowse, K. R., A. A. Avilion and C. W. Greider (1993). "Identification of a nonprocessive telomerase activity from mouse cells." Proc Natl Acad Sci U S A **90**(4): 1493-1497.
Puebla-Osorio, N., D. B. Lacey, F. W. Alt and C. Zhu (2006). "Early embryonic lethality due to targeted inactivation of DNA ligase III." Mol Cell Biol **26**(10): 3935-3941.
Qiao, F. and T. R. Cech (2008). "Triple-helix structure in telomerase RNA contributes to catalysis." Nat Struct Mol Biol **15**(6): 634-640.
Quinlan, A. R. and I. M. Hall (2010). "BEDTools: a flexible suite of utilities for comparing genomic features." Bioinformatics **26**(6): 841-842.
Rahman, N. and M. R. Stratton (1998). "The genetics of breast cancer susceptibility." Annu Rev Genet **32**: 95-121.
Rai, R., Y. Chen, M. Lei and S. Chang (2016). "TRF2-RAP1 is required to protect telomeres from engaging in homologous recombination-mediated deletions and fusions." Nat Commun **7**: 10881.

Rass, E., A. Grabarz, I. Plo, J. Gautier, P. Bertrand and B. S. Lopez (2009). "Role of Mre11 in chromosomal nonhomologous end joining in mammalian cells." Nat Struct Mol Biol **16**(8): 819-824.
Reddel, R. R. (2014). "Telomere maintenance mechanisms in cancer: clinical implications." Curr Pharm Des **20**(41): 6361-6374.
Reid, D. A., S. Keegan, A. Leo-Macias, G. Watanabe, N. T. Strande, H. H. Chang, B. A. Oksuz, D. Fenyo, M. R. Lieber, D. A. Ramsden and E. Rothenberg (2015). "Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair." Proc Natl Acad Sci U S A **112**(20): E2575-2584.
Riha, K., M. L. Heacock and D. E. Shippen (2006). "The role of the nonhomologous end-joining DNA double-strand break repair pathway in telomere biology." Annu Rev Genet **40**: 237-277.
Rijkers, T., J. Van Den Ouweland, B. Morolli, A. G. Rolink, W. M. Baarends, P. P. Van Sloun, P. H. Lohman and A. Pastink (1998). "Targeted inactivation of mouse RAD52 reduces homologous recombination but not resistance to ionizing radiation." Mol Cell Biol **18**(11): 6423-6429.
Ritchie, K., C. Seah, J. Moulin, C. Isaac, F. Dick and N. G. Berube (2008). "Loss of ATRX leads to chromosome cohesion and congression defects." J Cell Biol **180**(2): 315-324.
Roger, L., R. E. Jones, N. H. Heppel, G. T. Williams, J. R. Sampson and D. M. Baird (2013). "Extensive telomere erosion in the initiation of colorectal adenomas and its association with chromosomal instability." J Natl Cancer Inst **105**(16): 1202-1211.
Ropars, V., P. Drevet, P. Legrand, S. Baconnais, et al. (2011). "Structural characterization of filaments formed by human Xrcc4-Cernunnos/XLF complex involved in nonhomologous DNA end-joining." Proc Natl Acad Sci U S A **108**(31): 12663-12668.
Rosen, E. M. and M. J. Pishvaian (2014). "Targeting the BRCA1/2 tumor suppressors." Curr Drug Targets **15**(1): 17-31.
Rosenbloom, K. R., J. Armstrong, G. P. Barber, J. Casper, et al. (2015). "The UCSC Genome Browser database: 2015 update." Nucleic Acids Res **43**(Database issue): D670-681.
Roth, D. B. and J. H. Wilson (1986). "Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction." Mol Cell Biol **6**(12): 4295-4304.
Roy, S., S. N. Andres, A. Vergnes, J. A. Neal, Y. Xu, Y. Yu, S. P. Lees-Miller, M. Junop, M. Modesti and K. Meek (2012). "XRCC4's interaction with XLF is required for coding (but not signal) end joining." Nucleic Acids Res **40**(4): 1684-1694.
Ruhanen, H., K. Ushakov and T. Yasukawa (2011). "Involvement of DNA ligase III and ribonuclease H1 in mitochondrial DNA replication in cultured human cells." Biochimica et biophysica acta **1813**(12): 2000-2007.
Ruis, B. L., K. R. Fattah and E. A. Hendrickson (2008). "The catalytic subunit of DNA-dependent protein kinase regulates proliferation, telomere length, and genomic stability in human somatic cells." Mol Cell Biol **28**(20): 6182-6195.
Salvati, E., M. Scarsella, M. Porru, A. Rizzo, et al. (2010). "PARP1 is activated at telomeres upon G4 stabilization: possible target for telomere-based therapy." Oncogene **29**(47): 6280-6293.

Samper, E., F. A. Goytisolo, J. Menissier-de Murcia, E. Gonzalez-Suarez, J. C. Cigudosa, G. de Murcia and M. A. Blasco (2001). "Normal telomere length and chromosomal end capping in poly(ADP-ribose) polymerase-deficient mice and primary cells despite increased chromosomal instability." J Cell Biol **154**(1): 49-60.

Samper, E., F. A. Goytisolo, P. Slijepcevic, P. P. van Buul and M. A. Blasco (2000). "Mammalian Ku86 protein prevents telomeric fusions independently of the length of TTAGGG repeats and the G-strand overhang." EMBO Rep **1**(3): 244-252.

San Filippo, J., P. Sung and H. Klein (2008). "Mechanism of eukaryotic homologous recombination." Annu Rev Biochem **77**: 229-257.

Sarai, N., W. Kagawa, T. Kinebuchi, A. Kagawa, K. Tanaka, K. Miyagawa, S. Ikawa, T. Shibata, H. Kurumizaka and S. Yokoyama (2006). "Stimulation of Dmc1-mediated DNA strand exchange by the human Rad54B protein." Nucleic Acids Res **34**(16): 4429-4437.

Satoh, M. S. and T. Lindahl (1992). "Role of poly(ADP-ribose) formation in DNA repair." Nature **356**(6367): 356-358.

Schreiber, V., F. Dantzer, J. C. Ame and G. de Murcia (2006). "Poly(ADP-ribose): novel functions for an old molecule." Nat Rev Mol Cell Biol **7**(7): 517-528.

Schulte-Uentrop, L., R. A. El-Awady, L. Schliecker, H. Willers and J. Dahm-Daphi (2008). "Distinct roles of XRCC4 and Ku80 in non-homologous end-joining of endonuclease- and ionizing radiation-induced DNA double-strand breaks." Nucleic Acids Res **36**(8): 2561-2569.

Schwartzentruber, J., A. Korshunov, X. Y. Liu, D. T. Jones, et al. (2012). "Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma." Nature **482**(7384): 226-231.

Sebesta, M., P. Burkovics, S. Juhasz, S. Zhang, J. E. Szabo, M. Y. Lee, L. Haracska and L. Krejci (2013). "Role of PCNA and TLS polymerases in D-loop extension during homologous recombination in humans." DNA Repair (Amst) **12**(9): 691-698.

Sedivy, J. M. (1998). "Can ends justify the means?: telomeres and the mechanisms of replicative senescence and immortalization in mammalian cells." Proc Natl Acad Sci U S A **95**(16): 9078-9081.

Seluanov, A., D. Mittelman, O. M. Pereira-Smith, J. H. Wilson and V. Gorbunova (2004). "DNA end joining becomes less efficient and more error-prone during cellular senescence." Proc Natl Acad Sci U S A **101**(20): 7624-7629.

Seo, B., C. Kim, M. Hills, S. Sung, et al. (2015). "Telomere maintenance through recruitment of internal genomic regions." Nat Commun **6**: 8189-8199.

Sfeir, A. and T. de Lange (2012). "Removal of shelterin reveals the telomere end-protection problem." Science **336**(6081): 593-597.

Sfeir, A., S. T. Kosiyatrakul, D. Hockemeyer, S. L. MacRae, J. Karlseder, C. L. Schildkraut and T. de Lange (2009). "Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication." Cell **138**(1): 90-103.

Shahar, O. D., E. V. Raghu Ram, E. Shimshoni, S. Hareli, E. Meshorer and M. Goldberg (2012). "Live imaging of induced and controlled DNA double-strand break formation reveals extremely low repair by homologous recombination in human cells." Oncogene **31**(30): 3495-3504.

Shay, J. W., R. R. Reddel and W. E. Wright (2012). "Cancer. Cancer and telomeres--an ALTernative to telomerase." Science **336**(6087): 1388-1390.

Shinohara, A., H. Ogawa, Y. Matsuda, N. Ushio, K. Ikeo and T. Ogawa (1993). "Cloning of human, mouse and fission yeast recombination genes homologous to RAD51 and recA." Nat Genet **4**(3): 239-243.

Simsek, D., E. Brunet, S. Y. Wong, S. Katyal, et al. (2011). "DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation." PLoS Genet **7**(6): e1002080.

Simsek, D., A. Furda, Y. Gao, J. Artus, E. Brunet, A. K. Hadjantonakis, B. Van Houten, S. Shuman, P. J. McKinnon and M. Jasin (2011). "Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair." Nature **471**(7337): 245-248.

Simsek, D. and M. Jasin (2010). "Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4-ligase IV during chromosomal translocation formation." Nat Struct Mol Biol **17**(4): 410-416.

Simsek, D. and M. Jasin (2011). "DNA ligase III: a spotty presence in eukaryotes, but an essential function where tested." Cell Cycle **10**(21): 3636-3644.

Smith, F. W. and J. Feigon (1992). "Quadruplex structure of Oxytricha telomeric DNA oligonucleotides." Nature **356**(6365): 164-168.

Sprung, C. N., T. M. Bryan, R. R. Reddel and J. P. Murnane (1997). "Normal telomere maintenance in immortal ataxia telangiectasia cell lines." Mutat Res **379**(2): 177-184.

Stagno D'Alcontres, M., A. Mendez-Bermudez, J. L. Foxon, N. J. Royle and P. Salomoni (2007). "Lack of TRF2 in ALT cells causes PML-dependent p53 activation and loss of telomeric DNA." J Cell Biol **179**(5): 855-867.

Stansel, R. M., T. de Lange and J. D. Griffith (2001). "T-loop assembly in vitro involves binding of TRF2 near the 3' telomeric overhang." EMBO J **20**(19): 5532-5540.

Strom, C. E., F. Johansson, M. Uhlen, C. A. Szigyarto, K. Erixon and T. Helleday (2011). "Poly (ADP-ribose) polymerase (PARP) is not involved in base excision repair but PARP inhibition traps a single-strand intermediate." Nucleic Acids Res **39**(8): 3166-3175.

Suram, A., J. Kaplunov, P. L. Patel, H. Ruan, et al. (2012). "Oncogene-induced telomere dysfunction enforces cellular senescence in human cancer precursor lesions." EMBO J **31**(13): 2839-2851.

Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein and F. W. Stahl (1983). "The double-strand-break repair model for recombination." Cell **33**(1): 25-35.

Tebbs, R. S., L. H. Thompson and J. E. Cleaver (2003). "Rescue of Xrcc1 knockout mouse embryo lethality by transgene-complementation." DNA Repair (Amst) **2**(12): 1405-1417.

Tsai, A. G., H. Lu, S. C. Raghavan, M. Muschen, C. L. Hsieh and M. R. Lieber (2008). "Human chromosomal translocations at CpG sites and a theoretical basis for their lineage and stage specificity." Cell **135**(6): 1130-1142.

Tutt, A., M. Robson, J. E. Garber, S. M. Domchek, et al. (2010). "Oral poly(ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial." Lancet **376**(9737): 235-244.

Ulaner, G. A., H. Y. Huang, J. Otero, Z. Zhao, et al. (2003). "Absence of a telomere maintenance mechanism as a favorable prognostic factor in patients with osteosarcoma." Cancer Res **63**(8): 1759-1763.

van Overbeek, M. and T. de Lange (2006). "Apollo, an Artemis-related nuclease, interacts with TRF2 and protects human telomeres in S phase." Curr Biol **16**(13): 1295-1302.
Verdun, R. E. and J. Karlseder (2007). "Replication and protection of telomeres." Nature **447**(7147): 924-931.
Verkaik, N. S., R. E. Esveldt-van Lange, D. van Heemst, H. T. Bruggenwirth, J. H. Hoeijmakers, M. Z. Zdzienicka and D. C. van Gent (2002). "Different types of V(D)J recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells." Eur J Immunol **32**(3): 701-709.
Vilenchik, M. M. and A. G. Knudson (2003). "Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer." Proc Natl Acad Sci U S A **100**(22): 12871-12876.
von Kobbe, C., J. A. Harrigan, A. May, P. L. Opresko, L. Dawut, W. H. Cheng and V. A. Bohr (2003). "Central role for the werner syndrome protein/poly(ADP-ribose) polymerase 1 complex in the poly(ADP-ribosyl)ation pathway after DNA damage." Mol Cell Biol **23**(23): 8601-8613.
Walker, J. R., R. A. Corpina and J. Goldberg (2001). "Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair." Nature **412**(6847): 607-614.
Wang, F., E. R. Podell, A. J. Zaug, Y. Yang, P. Baciu, T. R. Cech and M. Lei (2007). "The POT1-TPP1 telomere complex is a telomerase processivity factor." Nature **445**(7127): 506-510.
Wang, H., B. Rosidi, R. Perrault, M. Wang, L. Zhang, F. Windhofer and G. Iliakis (2005). "DNA ligase III as a candidate component of backup pathways of nonhomologous end joining." Cancer Res **65**(10): 4020-4030.
Wang, H., Z. C. Zeng, T. A. Bui, E. Sonoda, M. Takata, S. Takeda and G. Iliakis (2001). "Efficient rejoining of radiation-induced DNA double-strand breaks in vertebrate cells deficient in genes of the RAD52 epistasis group." Oncogene **20**(18): 2212-2224.
Wang, M., W. Wu, W. Wu, B. Rosidi, L. Zhang, H. Wang and G. Iliakis (2006). "PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways." Nucleic Acids Res **34**(21): 6170-6182.
Wang, R. C., A. Smogorzewska and T. de Lange (2004). "Homologous recombination generates T-loop-sized deletions at human telomeres." Cell **119**(3): 355-368.
Wang, X., L. Liu, C. Montagna, T. Ried and C. X. Deng (2007). "Haploinsufficiency of Parp1 accelerates Brca1-associated centrosome amplification, telomere shortening, genetic instability, apoptosis, and embryonic lethality." Cell Death Differ **14**(5): 924-931.
Wang, Y., G. Ghosh and E. A. Hendrickson (2009). "Ku86 represses lethal telomere deletion events in human somatic cells." Proc Natl Acad Sci U S A **106**(30): 12430-12435.
Wang, Z. Q., B. Auer, L. Stingl, H. Berghammer, D. Haidacher, M. Schweiger and E. F. Wagner (1995). "Mice lacking ADPRT and poly(ADP-ribosyl)ation develop normally but are susceptible to skin disease." Genes Dev **9**(5): 509-520.
Wei, C., R. Skopp, M. Takata, S. Takeda and C. M. Price (2002). "Effects of double-strand break repair proteins on vertebrate telomere structure." Nucleic Acids Res **30**(13): 2862-2870.

Weltner, J., A. Anisimov, K. Alitalo, T. Otonkoski and R. Trokovic (2012). "Induced pluripotent stem cell clones reprogrammed via recombinant adeno-associated virus-mediated transduction contain integrated vector sequences." J Virol **86**(8): 4463-4467.

Wesoly, J., S. Agarwal, S. Sigurdsson, W. Bussen, et al. (2006). "Differential contributions of mammalian Rad54 paralogs to recombination, DNA damage repair, and meiosis." Mol Cell Biol **26**(3): 976-989.

West, R. B., M. Yaneva and M. R. Lieber (1998). "Productive and nonproductive complexes of Ku and DNA-dependent protein kinase at DNA termini." Mol Cell Biol **18**(10): 5908-5920.

Weterings, E., N. S. Verkaik, G. Keijzers, B. I. Florea, S. Y. Wang, L. G. Ortega, N. Uematsu, D. J. Chen and D. C. van Gent (2009). "The Ku80 carboxy terminus stimulates joining and artemis-mediated processing of DNA ends." Mol Cell Biol **29**(5): 1134-1142.

Windhofer, F., W. Wu and G. Iliakis (2007). "Low levels of DNA ligases III and IV sufficient for effective NHEJ." J Cell Phys **213**(2): 475-483.

Wong, L. H., J. D. McGhie, M. Sim, M. A. Anderson, S. Ahn, R. D. Hannan, A. J. George, K. A. Morgan, J. R. Mann and K. H. Choo (2010). "ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells." Genome Res **20**(3): 351-360.

Wu, P., M. van Overbeek, S. Rooney and T. de Lange (2010). "Apollo contributes to G overhang maintenance and protects leading-end telomeres." Mol Cell **39**(4): 606-617.

Xie, A., A. Kwok and R. Scully (2009). "Role of mammalian Mre11 in classical and alternative nonhomologous end joining." Nat Struct Mol Biol **16**(8): 814-818.

Xue, Y., R. Gibbons, Z. Yan, D. Yang, T. L. McDowell, S. Sechi, J. Qin, S. Zhou, D. Higgs and W. Wang (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): 10635-10640.

Yano, K., K. Morotomi-Yano, K. J. Lee and D. J. Chen (2011). "Functional significance of the interaction with Ku in DNA double-strand break recognition of XLF." FEBS Lett **585**(6): 841-846.

Ye, J., C. Lenain, S. Bauwens, A. Rizzo, et al. (2010). "TRF2 and apollo cooperate with topoisomerase 2alpha to protect human telomeres from replicative damage." Cell **142**(2): 230-242.

Yeager, T. R., A. A. Neumann, A. Englezou, L. I. Huschtscha, J. R. Noble and R. R. Reddel (1999). "Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body." Cancer Res **59**(17): 4175-4179.

You, Z. and J. M. Bailis (2010). "DNA damage and decisions: CtIP coordinates DNA repair and cell cycle checkpoints." Trends Cell Biology **20**(7): 402-409.

Zha, S., C. Boboila and F. W. Alt (2009). "Mre11: roles in DNA repair beyond homologous recombination." Nature Struct Mol Biol **16**(8): 798-800.

Zhang, Y. and M. Jasin (2011). "An essential role for CtIP in chromosomal translocation formation through an alternative end-joining pathway." Nat Struct Mol Biol **18**(1): 80-84.

Zheng, L. and B. Shen (2011). "Okazaki fragment maturation: nucleases take centre stage." J Mol Cell Biol **3**(1): 23-30.

Zhou, W. and P. W. Doetsch (1993). "Effects of abasic sites and DNA single-strand breaks on prokaryotic RNA polymerases." Proc Natl Acad Sci U S A **90**(14): 6601-6605.
Zhuang, J., G. Jiang, H. Willers and F. Xia (2009). "Exonuclease function of human Mre11 promotes deletional nonhomologous end joining." The Journal of biological chemistry **284**(44): 30565-30573.