

**Interaction Between The Three DNA  
Double Strand Break Repair Pathways in Human Cells**

**A DISSERTATION  
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF MINNESOTA  
BY**

**Sehyun Oh**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY**

**Eric A. Hendrickson, PhD  
Advisor**

**December 2011**



## Acknowledgements

First and foremost, I would like to express sincere gratitude to my adviser, Dr. Eric Hendrickson for his support through out my graduate study. When I struggled for my work and life, he gave me valuable advices and encouragements. They have been tremendous help and allow me to be a better scientist. Also, I am truly thankful of his unlimited patience and trust.

I would like to thank all the past and current members of Hendrickson's laboratory. My special thanks go to Jacob M. Zimbric and Dr. Yongbao Wang for all their selfless assistance and stimulating discussion through our collaboration. Also, I feel grateful for Eu Han Lee, Pauli Jackson, Matthew Mueller, Adam Harvey, Yinan Kan, Brian Ruis about their supports, helpful discussions and friendship that brought pleasure of working.

I would also like to thank the members of my thesis committee, Dr. Dennis Livingston, Dr. Alan Hooper, Dr. David Kirkpatrick and Dr. Nikunj Somia for their critical reviews of my scientific progress and their helpful suggestions.

I also express thanks to the past and current members of Bielinsky, Livingston and Harris Laboratories. I sincerely appreciate all the help from Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota.

Finally, I would like to show my appreciation to my family, especially my parents, my brother and my sister, for their endless support and encouragement during the difficult times of my graduate study. Their hearts and loves were always with me and made my life meaningful. I also want to extend thanks to my boyfriend and dance partner, Peter Westlake for his support and encouragement.

## **Dedication**

**To my parents**



## Abstract

Three DNA DSB repair pathways have been identified to date in human somatic cells: HR (homologous recombination), C-NHEJ (classic non-homologous end joining) and A-NHEJ (alternative-NHEJ). Which repair pathway gets used appears to be a function of several parameters although cell type and cell cycle phase are chief amongst them. For example, B- and T-lymphocytes carry out site-specific recombination {V(D)J and class switch} reactions that are required for the proper development of the human immune system. Since V(D)J and class switch recombination reaction rely heavily on C-NHEJ to produce functional immune molecules, the C-NHEJ repair process dominates in B-and T-lymphocytes. Similarly, cells that reside in the G1/G0 phase of the cell cycle use C-NHEJ almost exclusively and since ~75% of the cells in a human's body are, at any given time, in G1/G0, the C-NHEJ repair pathway is again chiefly used.

Seven core proteins; Ku70, Ku86, DNA-PK<sub>cs</sub>, Artemis, LIGIV, XRCC4 and XLF are known to be required for C-NHEJ. Mutations of most of these genes result in cells or animals with profound deficits in DNA DSB repair, hypersensitivity to DNA damaging agent exposure and severe immune deficiencies. The phenotypes resulting from loss of function of LIGIV, were, however rather variable depending upon what model organism was being study. Consequently, it was unclear whether LIGIV is essential for human cell survival or not. To experimentally address this issue, we functionally inactivated the LIGIV gene and showed that LIGIV (and presumably C-NHEJ) is not essential for human somatic cell survival.

LIGIV is but one of three mammalian ligases, the other two being LIGI and LIGIII. LIGI is thought to be involved in HR and DNA replication. LIGIII was recently implicated as possibly being involved in A-NHEJ pathway. There are, however, to date very few model systems available to study A-NHEJ. Recent studies carried out in mice indicated that LIGIII had an essential function in mitochondria and was likely involved in chromosomal translocations in nucleus. To study similar phenotypes in human somatic cells we generated a conditionally viable LIGIII-null cell line. We then demonstrated that human LIGIII is also essential for mitochondria function but then went on to show that absence of nuclear LIGIII is tolerated with almost no pathological phenotypes.

Finally, we have observed that cells lacking the key C-NHEJ factor, Ku86, succumb because of telomere dysfunction, including telomere loss, sister chromatid fusion and t-circle formation. Several laboratories have postulated that the t-circle formation is due to the aberrant activity of HR on telomeric DNA. To experimentally test this model, we functionally inactivated the Rad54B gene — a major HR component — in Ku86 conditional null cells. Impressively, a Rad54B deficiency partially rescues the defects of Ku86-null cells: it significantly delays the death of Ku86-null cells and significantly reduces t-circle formation. These results indicate that the telomere dysfunction observed in Ku86-null cell is, at least in part, mediated by HR.

# TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS .....</b>	<b>I</b>
<b>DEDICATION.....</b>	<b>II</b>
<b>ABSTRACT .....</b>	<b>III</b>
<b>TABLE OF CONTENTS .....</b>	<b>V</b>
<b>LIST OF TABLES .....</b>	<b>VII</b>
<b>LIST OF FIGURES .....</b>	<b>VIII</b>
<b>CHAPTER I: INTRODUCTION .....</b>	<b>1</b>
<b>CHAPTER II: LIGIV IS NOT ESSENTIAL FOR THE SURVIVAL OF HUMAN SOMATIC CELLS .....</b>	<b>34</b>
INTRODUCTION.....	36
MATERIALS AND METHODS.....	40
RESULTS.....	44
DISCUSSION.....	49
<b>CHAPTER III: LIGIII IS AN ESSENTIAL GENE FOR THE SURVIVAL OF HUMAN SOMATIC CELLS.....</b>	<b>72</b>
INTRODUCTION.....	74
MATERIALS AND METHODS.....	79
RESULTS.....	83
DISCUSSION.....	90
<b>CHAPTER IV: A DEFICIENCY IN RAD54B RESCUES THE TELOMERE DEFECTS AND LEADS TO THE PROLONGED SURVIVAL OF KU86-NULL CELLS.....</b>	<b>111</b>
INTRODUCTION.....	113
MATERIALS AND METHODS.....	117
RESULTS.....	120

DISCUSSION.....	125
<b>CHAPTER V: CONCLUSIONS AND FUTURE DIRECTIONS.....</b>	<b>142</b>
<b>BIBLIOGRAPHY .....</b>	<b>160</b>

## LIST OF TABLES

### **Chapter II: LIGIV is not essential for the survival of human somatic cells**

**Table 1.** Summary of GCR frequency in LIGIV clones ..... 68

**Supplementary Table 1.** Gene targeting rate in LIGIV deficient cells..... 71

### **Chapter IV: A deficiency in Rad54B rescues the telomere defects and leads to the prolonged survival of Ku86-null cells**

**Table 1.** Summary of Rad54B gene targeting rate in LIGIV-null and Ku86-conditional null cell lines..... 139

## LIST OF FIGURES

### **Chapter II: LIGIV is not essential for the survival of human somatic cells**

<b>Figure 1.</b> Targeting construct and primers used for LIGIV gene targeting .....	54
<b>Figure 2.</b> Immunoblotting analysis of LIGIV targeted clones .....	56
<b>Figure 3.</b> LIGIV deficiency causes mild growth retardation .....	57
<b>Figure 4.</b> LIGIV null cell displays extreme sensitivity against etoposide.....	60
<b>Figure 5.</b> Microhomology mediated DNA end-joining activity in LIGIV-deficient cells... .....	62
<b>Figure 6.</b> Immunoblot result of LIGIV complemented clones .....	64
<b>Figure 7.</b> LIGIV complementation rescues extreme etoposide-sensitivity of LIGIV null cell .....	66
<b>Supplementary Figure 1.</b> G-band karyotype of LIGIV deficient cells .....	69

### **Chapter III: LIGIII is an essential gene for the survival of human somatic cells**

<b>Figure 1.</b> Scheme for functional inactivation of the human LIGIII locus .....	95
<b>Figure 2.</b> Confirmation of DNA LIGIII conditional null cell line .....	97
<b>Figure 3.</b> Complementation with a mitochondrial-exclusive form of LIGIII rescues the cell lethality .....	99
<b>Figure 4.</b> LIGIII deficiency causes growth retardation .....	101
<b>Figure 5.</b> LIGIII-null cells do not show sensitivity to DNA damaging agents .....	103
<b>Figure 6.</b> LIGIII deficiency has no effect on the overall DNA end-joining ability of human somatic cells .....	105
<b>Figure 7.</b> LIGIII-null cells still retain microhomology-mediated repair activity.....	107
<b>Figure 8.</b> LIGIII deficiency has no effect on rAAV gene targeting rate.....	109

### **Chapter IV: A deficiency in Rad54B rescues the telomere defects and leads to the prolonged survival of Ku86-null cells**

<b>Figure 1.</b> The Rad54B targeting strategy.....	129
<b>Figure 2.</b> Western blot analysis confirm the construction of a Ku86:Rad54B doubly-null cell line.....	131
<b>Figure 3.</b> A 4-OHT-inducible CreERT2 system for use with the Ku86:Rad54B clones	

.....	133
<b>Figure 4.</b> A Rad54B deficiency attenuates the lethality of Ku86 loss of function.....	135
<b>Figure 5.</b> Neutral/neutral 2D-gel electrophoresis confirms reduced t-circle formation .....	137
<b>Supplementary Figure 1.</b> Targeting procedure for triploid Rad54B gene.....	141

# **CHAPTER I**

## **INTRODUCTION**



Living organisms store and transmit their genetic information in the form of DNA, so the reliable replication and sustainment of DNA's integrity is important. Among the various types of DNA damage that a cell can sustain, DNA double-strand breaks (DSBs) are the most deleterious because even a single un-repaired DSB can cause cell death (1). Not surprisingly, therefore, the number of DSBs that a cell suffers is thought to be quite small, in the range of only ~10 DSB per day. It is wrong, however to view all DSBs as pathological (2). In mammals, for example, the generation of site-specific DSBs is a critical step during antigen receptor gene rearrangement in the course of lymphocyte maturation process, which includes V(D)J recombination and class-switch recombination (CSR) (3-5). In addition, DSB formation is an absolute prerequisite for proper generation of human gametes during meiosis (6). Nonetheless, it is probably fair to describe the majority of DSBs as harmful. Some pathological DSBs can be generated as the by-product of natural processes. First, DNA replication across a nick can convert that single-strand lesion into a DSB (7). Second, as a by-product of oxidation within mitochondria, cells generate high-energy, reactive oxygen species (ROS). ROS cause damage to mitochondrial genes and once it is leaked into the nucleus, ROS can generate DSBs in chromosomal DNA (2). In addition to the internal generation of DSBs, however, DSBs can also be produced from exogenous sources. The most common of these is due to ionizing radiation (IR) or chemotherapeutic exposure during a course of clinical cancer therapy (2). Additionally, DNA damage likely accrues due to exposure to common environmental toxins (*e.g.*, automobile exhaust, cigarette smoke, sunlight, *etc.*). Importantly, no matter whether the DSBs occur endogenously or exogenously, their impact is almost always deleterious to the well-being of the cell.

Unrepaired DSB can lead to the accumulation of mutations, cancer and cell death. To protect individuals and their progeny from the detrimental results of DSBs, all cells have evolved several different ways to repair DSBs.

### **DNA DSB repair pathways**

There are two major DNA DSB repair pathways; homologous recombination (HR) and classic non-homologous end joining (C-NHEJ) (2, 8-12). They are evolutionarily conserved from bacteria to man but their usage varies depending on the organism at-hand

and the cell cycle. HR, as the name implies, requires extensive sequence homology to template. In lower eukaryotes like yeast, HR is the predominant repair mechanism (10, 11, 13-16). In the higher eukaryotes, however, HR is the major repair mechanism only during the late S/G2 phases of the cell cycle, when proximal homology donors are present as sister chromatids or homologous chromosome. Because HR uses homologous sequences as a repair template, the repaired products are generally regarded as “error-free”. This descriptor is essentially accurate for most HR-mediated repair events although it should be kept in mind that HR is likely the cause of rarer mutagenic loss-of-heterozygosity events (17). The other major DNA DSB repair pathway is C-NHEJ, which — again as the name implies — does not require a long homologous sequence donor, but instead simply re-joins the ends together in a process that often results in the loss or addition of multiple nucleotides (11, 15, 18, 19). Consequently, C-NHEJ is often (correctly) referred to as an error-prone process (20). Despite this mutagenic attribute, however, C-NHEJ is nonetheless the main DSB repair pathway in higher organisms, — especially in humans — and is most active during the G1/G0 phases of the cell cycle (16, 21, 22). The fact that the vast majority of the human genome is littered with non-coding DNA probably explains why C-NHEJ is tolerated so well, since minor alterations to non-coding regions likely have little or no evolutionary impact. Recently, a third, minor DSB repair pathway has been reported. This novel pathway, alternatively named either alternative non-homologous end joining (A-NHEJ) or backup non-homologous end joining (B-NHEJ), is detectable only under C-NHEJ-deficient conditions (23-27). A-NHEJ requires 5 to 25 nucleotides of micro-homology, and its repair products always contain deletions. Conceptually, A-NHEJ can be thought of as a hybrid of the HR and C-NHEJ pathways since it requires (micro) homology (*i.e.*, ~HR) to mediate direct end-joining in a mutagenic fashion (*i.e.*, ~NHEJ). The details of the A-NHEJ pathway, however, are still vague and its detailed mechanism is under active investigation.

### **DSB signaling**

The first step in DNA DSB repair, regardless of which DNA DSB repair is ultimately utilized, is likely to be the upstream recognition of the DSB and the signaling that ensues to activate the downstream repair pathways. Like many biological processes, DSB

signaling is mediated by a kinase-mediated phosphorylation cascade (28, 29). There are two related kinases required for the DNA DSB damage response: ATM (ataxia-telangiectasia mutated) and ATR (ATM- and Rad3- related). They are members of the PIKK (phosphatidylinositol 3-kinase-like kinase) family (30, 31). Both ATM and ATR are the principle DNA damage signal transducers but they deal with different types of substrates; ATM mostly functions in DSB repair whereas ATR responds more to ssDNA and consequently more to the DSBs introduced at stalled DNA replication forks (29, 32-34). ATM and ATR are recruited to damage sites through the interaction with Nbs1 (part of the MRN complex; see below) and ATR-interacting protein (ATRIP), respectively (35). Damage sensing subsequently proceeds via a checkpoint signaling cascade, which includes phosphorylation of downstream effectors, chromatin remodeling, recruitment and retention of additional proteins required for repair process (36).

ATM is activated via a two-step mechanism (37). ATM is normally present as an inactive dimer, but following the introduction of DNA damage, ATM is recruited to the damaged sites through the interaction with Nbs1. This step results in MRN-dependent DNA-tethering, which increases the local concentration of dsDNA ends and enhances the monomer formation (37). Second, the ATM monomer undergoes *trans* auto-phosphorylation on Ser1981, which happens in a DNA-independent manner and fully activates ATM (38, 39). Fully-activated ATM subsequently phosphorylates (and thus activates) many downstream cellular targets such as Chk2 (checkpoint kinase 2), p53, H2AX and BRCA1 (breast cancer allele 1) to complete the signaling process (8).

ATM signaling is required for the bulk of DSB signaling. The DSBs associated with replication fork collapse, however, are a special case where ATR is the most relevant kinase. During normal DNA replication, the single-stranded DNA that accumulates behind a replication fork on the retrograde strand is normally coated with RPA (replication protein A), a trimeric ssDNA binding complex. If the replication fork should collapse and generate a DSB, ATRIP (ATR-interacting protein) is recruited to the RPA-bound ssDNA (35, 40, 41). ATRIP, in turn, recruits ATR, which becomes activated and phosphorylates its targets, including RPA and Chk1 (checkpoint kinase 1), to initiate checkpoint response (38, 40).

Initially, it was thought that ATM and ATR work independently of each other. However, more recent evidence suggests that these two signaling pathways may actually overlap more than was previously appreciated. For example, following exposure to IR in S/G2 phases, ATM activates the nuclease activity of the MRN complex, which generates RPA-coated ssDNA that becomes a substrate for ATR (34). Similarly, after DNA replication fork stalling the phosphorylation and activation of ATM is ATR-dependent (42). In summary, ATM and ATR, either independently or together, seem to regulate cellular responses to DSBs.

### **MRN complex**

As detailed above, the recruitment of ATM to DSBs requires interaction with the protein Nbs1, which is one component of the hetero-hexameric MRN protein complex that is composed of two molecules of Mre11, Rad50, and Nbs1 (Xrs2 in yeast). MRN is a key, but very complicated complex, that is required for both upstream regulatory and downstream processing steps. In addition to the activation of ATM, the rapid recruitment of MRN to a DSB that will be repaired by HR ensures the protection of the broken ends from non-specific degradation and it keeps the broken ends in proximity to one another (43). MRN, which has associated helicase and nuclease activities, also functions in the end-processing steps of A-NHEJ. In contrast to HR and A-NHEJ, C-NHEJ confusingly appears to only utilize the MRN complex for damage signaling and not for end resection (44). Interestingly, the MRN complex is also found at telomeres and seems to facilitate t-loop formation by generating 3' G-overhangs using its nuclease activity (discussed in more detail below) (43, 45-48). Despite intense effort, it is still unclear how — depending on the particular repair pathway utilized — the end-processing activities of MRN complex are regulated (44). At least some of these complicated attributes of MRN appear to be coordinated by different combinations of conformational changes, post-translational modifications, and protein interactions (43, 49).

Mre11 has both Rad50 and Nbs1 binding domains in addition to a DBD (DNA binding domain) (49-51). Mre11 is thus likely through these three domains to be the linchpin that mediates the interactions of the individual subunits with each other and with DNA (51). Mre11 also has both endonuclease and 3' to 5' exonuclease activities. While

Mre11 is required for end processing, it is not clear whether MRN actually directly does this, because Mre11 does not have the 5' to 3' polarity exonuclease activity generally necessary for end resection (52-54). Rad50 has globular heads containing ATPase domains at both ends and a coiled-coil helix with a zinc hook at the middle (50). The ATPase and zinc hook domains are required for ATP-dependent DNA unwinding activity and DNA end tethering, respectively (55). The third subunit, Nbs1, has two BRCT (breast cancer allele C-terminus) domains that can interact with ATM or ATR, when MRN complex is recruited to DSBs or stalled replication forks, respectively. ATM directly interacts with Nbs1 and the MRN-dependent accumulation of ATM at DSBs occurs within seconds (39). Nbs1 autoregulates the catalytic functions of the MRN complex by activating ATM, which ultimately results in ATM or ATR subsequently phosphorylating Nbs1. Nbs1 is required for localization of MRN complex in the nucleus as well (56, 57).

Defects in the MRN complex cause serious diseases. Nbs1-deficiency leads to a condition called Nijmegen breakage syndrome (NBS) (58). Patients without functional Mre11 have ataxia-telangiectasia-like disorder (ATLD). Only one Rad50 patient, who displayed NBS-like phenotypes, has been reported (59). All MRN patients still express mutant forms of their respective protein, because complete loss of MRN does not appear to be tolerated in mammals (60).

### **Homologous recombination (HR)**

HR is an evolutionarily well-conserved DSB repair pathway, which uses homologous regions of either a sister chromatid or a homologous chromosome as a template for repair (10, 13). HR, while perhaps not the major DNA DSB repair pathway, is nonetheless essential. HR is required for all the meiotic recombination that occurs during the generation of functional gametes (6, 61). Moreover, HR plays a crucial role in the restart and repair of stalled or broken replication forks during S phase (62). Lastly, HR also plays a role in the repair of spontaneous and environmentally-induced DNA damage that occurs to cells in late S/G2 phases of the cell cycle.

Although some of the specifics of HR are still lacking, the general process has been described in reasonable detail. A general overview of the process is as follows: a DSB is

recognized by damage sensors and processed by nucleases to yield 3'-ssDNA (single-stranded DNA). Several HR proteins then bind to and stabilize the ssDNA. The protein-coated ssDNA is then used in a search for homology and forms a synapsis with a homology target through strand-invasion. Replication starts from the synaptic point, where it uses the 3'-end of invading strand as a primer and an intact homologous sequence as a template. This replication results in the recombination junction migrating outward from the initial break site. This junction is ultimately resolved by resolvases and the remaining gaps are filled in by polymerases.

A large number of proteins are involved in mammalian HR, including Rad51, Rad51 paralogs (XRCC2, XRCC3, Rad51B, Rad51C, Rad51D), Rad52, the BRCA proteins, Rad54 & Rad54B, and the MRN complex. Rad51 and the Rad51 paralogs mediate strand-exchange reactions during HR. Rad51 paralogs, a family of five proteins, share only 20% to 30% amino acid homology with Rad51, but appear to have retained the strand transfer activity of the parent gene (63, 64). Through evolution, the Rad51 paralogs have presumably acquired novel functions, but the specific roles and tissue distributions of each member are not clear yet.

Rad52 does not have demonstrable enzymatic activity of any kind and yet it is required to facilitate the RAD51-dependent strand transfers and to help displace RPA from ssDNA (65). In yeast, Rad52 is arguably the most important HR protein and Rad52 mutants are essentially incapable of performing just about any HR-related reaction (66, 67). Confusingly, this importance is not conserved as mutation of Rad52 in the mouse (68) or human somatic cells {Yinan Kan, personnel communication} are virtually aphenotypic. It appears that in mammals Rad52 has evolved to work not in the canonical HR pathway but predominately in an HR subpathway termed single-strand annealing (69). This fact begs the question, however, about what is the cofactor for strand exchange in mammals if Rad52 has evolved different uses? In this regard, BRCA1 and BRCA2 are attractive candidates. Mutations in BRCA1 or BRCA2 increase the risk of developing breast, ovarian and other cancers and are thus true tumor suppressors (70). BRCA1 has broad cellular roles including DNA repair, chromatin remodeling, cell cycle regulation, and transcriptional regulation (71). Whether BRCA1 can stimulate Rad51 strand exchanges is currently unknown. In contrast, BRCA2 is clearly involved in DNA

recombination and repair, and facilitates Rad51-nucleoprotein filament formation (72, 73).

Rad54 is a dsDNA-dependent ATPase, which translocates along chromosomes as a chromatin remodeling protein. Rad54 is involved in several steps of HR (74). At the early phase of HR, Rad54 interacts with Rad51 and stabilizes the Rad51-nucleoprotein filament (75-77). Rad54 also promotes homology pairing and strand exchange steps (78-80). Later, Rad54 induces the dissociation of Rad51 from the repaired chromosome. Rad54B, a homolog of Rad54, shares many biochemical properties with Rad54 (81). Rad54B is also a chromatin-remodeling factor with DNA-dependent ATPase and helicase activities (82). Thus, Rad54B also has the ability to translocate down DNA and promote DNA helix opening. Moreover, Rad54B also interacts with Rad51 and facilitates homology pairing and D-loop formation (81). Interestingly, Rad54B has longer N-terminus, which has not yet been fully characterized, than RAD54 and this undoubtedly confers additional activities on to RAD54B that are not found in RAD54. Rad54B knockout somatic human cell lines are viable and they contain reduced HR activity, as indicated by their inability to carry out any (subsequent) gene targeting (83). Accentuating the difference between murine and human cell models, Rad54B knockout MEF cell lines display different phenotypes: the gene targeting rate is not reduced but the cells are hypersensitive to IR and mitomycin C, a DNA cross-linking agent (74). It is not understood why a Rad54B-deficiency in human and mouse cells results in different phenotypes, but this species difference is not that unusual (see, for more examples, below).

The MRN complex was introduced earlier as a trimeric complex, which is required for upstream DNA damage signaling. The exact role(s) of MRN, however, are complicated and it is clear that the complex also has direct roles in the downstream aspects of HR. As detailed above, the MRN complex is one of the first components recruited to DSBs and the DNA-bound MRN complex recruits ATM to the break through a direct interaction with the Nbs1 subunit (35, 39, 51, 84-86). Phosphorylation of MRN by ATM enhances the nuclease activity of the MRN complex, which is necessary for the end processing that generates the 3'-hydroxyl overhang required for all subsequent homology search and strand invasions (34, 60). Confusingly, however, the actual

exonuclease involved in 3' overhang generation is as yet unidentified, although it is almost certainly not MRN as Mre11 has 3' to 5' exonuclease activity, which is the opposite polarity required to generate a 3' overhang. The most likely candidate is CtIP (C-terminal interacting protein), which directly interacts with MRN complex through Nbs1 and seems to facilitate ssDNA formation at the site of DSB {there have been 3 or 4 papers in the past month showing that these MRN and CtIP work together to mediate end-processing; you should cite them}. It is possible that the interaction with CtIP mediates the transition of the MRN complex's activity from end sensing to end processing by switching the direction of or altering the exonuclease activity (22, 49, 87-89). Nonetheless, it is still not clear whether Mre11, CtIP, or both are required for end resection. Moreover, it has not even been demonstrated yet whether CtIP itself has an exonuclease activity or not. What seems clear (even if the mechanism is obscure) is that MRN and CtIP together are responsible for removing the first few nucleotides (and probably no more than a hundred) around the break. Subsequently, Exo1 (exonuclease 1) and Dna2 take over the end resection and generate upwards of kilobase-sized 3' ssDNA overhang (54, 90). Thus, the current model envisions resection as a two-step process (37): MRN/CtIP does the small initial resection, and Exo1/Dna2 performs the large subsequent resection. The biochemical mechanism of end resection is one of the most intensely studied areas in HR.

After end resection, the 3' overhang is covered with single-stranded DNA binding protein, RPA, which protects the ssDNA from secondary structure formation and nuclease attacks (91). Before RPA gets saturated on ssDNA (92, 93) and inhibits strand-exchange, Rad52 replaces RPA with Rad51 and stabilizes the Rad51:nucleoprotein filament (65). Rad52 also directly binds to ssDNA protecting it from nucleolytic degradation and facilitates strand-annealing (66). Homology searches by the Rad51-nucleoprotein filament are performed by random collision with duplex DNA (94). Once a 3' overhang finds a duplex homology target, strand invasion and joint molecule formation transpire followed by DNA replication, where the all-important 3'-hydroxyl end of invading strand is used as a replication primer. When this end of the newly synthesized strand is linked to the other side of the broken end, a Holliday junction is formed (95). A classic Holliday junction has four strands of DNA, which are derived



from two duplex DNA molecules and which are crossed, relative to each other. Specific resolvases are then needed to resolve this topologically complicated structure back into two dsDNAs. Holliday junction resolution appears to be carried out by at least 3 complexes (96-99), that are genetically and biochemically redundant with one another. In mitotic cells, the resolution step is generally not associated with crossing-over (*i.e.*, rather it is a gene conversion event) whereas during meiosis a large fraction of the events do trigger crossing-over, resulting in genetically recombinant chromosomes.

In summary, HR is relatively well-understood at the mechanistic level and in late S phase cells it can lead to error-free repair of DNA DSBs.

### **Classic non-homologous end joining (C-NHEJ)**

C-NHEJ is the major DNA DSB repair pathway in human somatic cells. In addition to DSB repair, C-NHEJ components also function in V(D)J recombination, CSR (class switch recombination), and telomere maintenance. V(D)J recombination is a site-specific DNA recombination process required for the maturation of antigen receptor genes in early B- and T-lymphocytes (100, 101). Lymphocyte-specific endonucleases, RAG1 (recombination activating gene 1) and RAG2, make breaks at specific recombination signal sequences (RSS) and generate two products: the signal joint and the (biologically important) coding joint. The proper generation of both of these products absolutely requires C-NHEJ (2). CSR is a temporally distinct recombination reaction following V(D)J recombination that replaces one antibody constant region with another, thus altering the antibody effector function (4). CSR is initiated by AID (activation-induced cytidine deaminase), which introduces uracil:guanine mismatches (102, 103). These mismatches are recognized by mismatch repair proteins or UNG (uracil-DNA glycosylase) and eventually converted into a DSB that can be repaired by C-NHEJ (104). Unlike V(D)J recombination intermediates, which are exclusively processed by C-NHEJ, CSR intermediates can be repaired by either C-NHEJ or, less frequently, A-NHEJ (alternative-NHEJ), a somewhat obscure DSB repair pathway, described more fully below. Lastly, telomeres are the nucleoprotein structures found at the end of linear chromosomes (47, 105). A telomere generally consists of a 2 to 15 kb stretch of repetitive DNA complexed with a dedicated set of proteins, which in humans are

collectively called “Shelterin” (106-108). Telomeres provide at least two functions for chromosomes: 1) they provide an elegant solution to the “End Replication Problem” (105, 109) and they stabilize the chromosomes from fusing with one another. It is this latter activity — that of impeding chromosome end fusions — that makes telomeres one of the least likely places to find DNA DSB factors, whose biological goal is almost always to join DNA fragments together. Paradoxically, however, many C-NHEJ factors actually play a positive role in telomere maintenance (110-119). Some of these requirements will be explained later in this chapter (120).

There are at least seven main C-NHEJ components: the Ku heterodimer (Ku70 and Ku86), DNA-PK<sub>cs</sub> (DNA dependent protein kinase catalytic subunit), Artemis, and a trimeric complex consisting of LIG4/XRCC4/XLF (ligase IV/X-ray cross complementing 4/XRCC4-like factor). In addition, DNA damage sensors, signal transducers, and polymerases are also required to complete DSB repair via C-NHEJ.

### **Ku70/86**

The Ku complex, a heterodimer composed of Ku70 and Ku86 proteins, was first identified in patients suffering from an autoimmune disease, polymyositis-scleroderma overlap syndrome (121). This clinical red herring (*i.e.*, a seeming autoimmune function) obscured for several decades Ku’s real role, which in reality turned out to be that of a critical DNA DSB repair factor. The first direct evidence that Ku is involved in DNA DSB repair and V(D)J recombination came when Chinese hamster ovary cell lines defective in DNA DSB repair and V(D)J recombination could be functionally rescued by the expression of a human Ku86 cDNA (122, 123). Ku is evolutionarily well conserved and homologs exist from bacteria to human. Ku70 and Ku86 (the names relate to the approximate sizes of the proteins, which are 69 and 83 kDa, respectively) together form a donut- or ring-shaped protein. Even when not bound to DNA, Ku70 and Ku86 form a heterodimer and stabilize each other. Consequently, mutations of or deficiencies in one of the subunits results in depletion of the other (124). Ku is a highly abundant (about 500,000 molecules per cell in humans; (125)) cellular protein and because the ring formed by the heterodimer is precisely big enough to encircle a dsDNA end, it has a very high affinity for binding to free DNA ends in a sequence-independent manner (126). In

addition to a sequence-independency, Ku complex can also bind to DSB ends with different structures such as blunt ends, ends with protrusions, or hairpin structures (127, 128). These features have been visualized by crystal structures of Ku complexed with dsDNA, where the interdigitated Ku70 and Ku86 form an opening that is precisely large enough to accompany only one DNA double strand (129). Once Ku binds to DNA, it displays a higher affinity for the other C-NHEJ components and appears to function as an assembly platform for them.

The central portion of the two Ku subunits forms the ring-shaped DNA end-binding domain (129). It is this central domain that is conserved throughout evolution (2, 116, 130). In contrast, the N- and C-termini of each subunit appear to have grown as the proteins evolved. Specifically, Ku interacts with a kinase (DNA-PK<sub>cs</sub>), a nuclease (Artemis), polymerases (pol  $\mu$  and  $\lambda$ ) and the ligase complex (LIGIV:XRCC4:XLF) and Ku makes these interactions using domains located within the extreme N- and C-termini of the two subunits. In vertebrates, one of the most important contacts (mediated by the Ku86 C-terminus) is with DNA-PK<sub>cs</sub> (131). This interaction is needed to recruit and activate the kinase at the DSB and is required for all of DNA-PK<sub>cs</sub>'s subsequent downstream signaling functions (131, 132). Ku makes similar contacts with most of the other C-NHEJ factors. Thus, polymerase  $\mu$  and  $\lambda$  contribute to junctional additions, which facilitate incompatible ends joining, and this activity requires their BRCT domains and is dependent both physically and functionally on Ku (133, 134). Similarly, several studies have demonstrated an interaction between Ku and LIGIV:XRCC4:XLF (135-138), and the requirement of Ku for the recruitment of the LIGIV:XRCC4:XLF complex to DNA ends (135). In summary, Ku is a multi-domain protein that uses its central domain to bind onto the ends of all ds DNA molecules and then uses its N- and C-terminal domains to recruit a bevy of other DNA repair factors to the end(s) to facilitate repair.

Quite surprisingly, Ku also has a role in telomere maintenance. In yeast, yKu80 directly interacts with telomeric DNA and this interaction is required for its telomere functions, such as telomere end protection, telomere length regulation, and sub-telomeric silencing (139). This observation supports the “two-face” model of Ku (140), where the yKu70:yKu80 heterodimer can bind DNA in an asymmetric orientation; when yKu80 is

distal from the break site the function of Ku complex is different from when yKu70 is proximal to the end. This model is further substantiated by a structure: function study showing that point mutations in yKu80 abrogate telomere functions, whereas yKu70 mutations affect DSB repair (140). Human Ku appears to have similar telomere functions as yeast Ku, but it is not clear yet whether human Ku interacts directly with telomeres or not. In human cells, Ku can be co-immunoprecipitated with the telomeric Shelterin complex and it can also pull down telomere DNA sequences in chromatin immunoprecipitation (ChIP) experiments (113). In these experiments, Ku seemed to interact with TRF1 (telomere recognition factor 1; an essential component of Shelterin), but not with the telomere directly (112). Additional studies demonstrated that Ku and TRF2 also interact (141). These studies support the belief that Ku is either a component of Shelterin or a critically-important associated factor. It has been suggested that telomere-bound Ku suppresses the pathological action of HR on telomeres and that Ku's interaction with the Shelterin complex prevents it from initiating C-NHEJ (142). Although Ku's exact role at human telomeres is not yet resolved, it is clear that Ku is essential, because Ku-deficient human somatic cells are dead due to massive telomere dysfunction (117). This essential requirement for Ku in telomere maintenance likely explains why not a single Ku-null patient has ever been reported.

In addition to Ku's well-documented role in DNA DSB repair and telomere maintenance, Ku has also been reported to function as a membrane protein in a subset of cells and contribute to adhesion and invasion processes (143-145). Membrane-bound Ku also affects viral or bacterial entry into target cells (146-148) and a requirement for Ku in cell-matrix attachment (149), proteolytic processing (150), and migration of activated monocytes (151) have been reported as well. In addition, cytoplasmic Ku70 also inhibits apoptosis by sequestering a pro-apoptotic protein, *Bax*, outside of mitochondria (152, 153). The biological significance of these non-nuclear roles for Ku is obscure since the loss-of-function of the nuclear activities of Ku are generally sufficient to explain all of the resultant phenotypes of Ku mutant cells.

Perhaps most puzzlingly, even though Ku is evolutionarily conserved, its importance shows considerable species variability. Humans, for example, are the only species where a Ku deficiency causes lethality. In contrast to humans, a variety of Ku-null model

organisms have been reported, such as in bacteria (154, 155), yeast (137), worms (156), plants (111, 118, 157), chickens (158, 159), hamster (123) and the mouse (160, 161). The less severe phenotypes observed in these “lower” organisms is probably linked to species differences in telomere maintenance mechanisms. For example, viable knockout mice for Ku70 and Ku86 have been reported (160-162). These mice have very severe deficits in DNA DSB repair and V(D)J recombination, but only mild telomere dysfunction. However, telomere maintenance in the mouse is fundamentally different than it is in humans. Mice have telomeres that are about 10 times longer than the average human telomere (163, 164). Moreover, all mouse somatic cells express the enzyme (telomerase) that can synthesize new telomeres, whereas human somatic cells do not (165). Thus, human cells appear to be much more sensitive to mutations that cause telomere dysfunctions than other species and this fact probably explains the differences in loss-of-function mutations for Ku reported in various species.

### **DNA-PK<sub>cs</sub> and Artemis**

DNA-PK<sub>cs</sub> and Artemis are not evolutionarily conserved and are present only in vertebrates (15, 166). The evolutionary introduction of two new proteins with the genesis of vertebrates clearly coincides with the development of V(D)J recombination (167). Importantly, V(D)J recombination, which is essential for establishing a functional immune system, results in the generation of an intermediate that requires unusual end processing. DNA-PK<sub>cs</sub> and Artemis seem to have evolved to deal with this processing event.

DNA-PK<sub>cs</sub>, a member of the PIKK family, is a serine/threonine kinase (30, 31). DNA-PK<sub>cs</sub> is a large protein with a molecular weight of 469 kDa, which is in itself perplexing since the only functionally identifiable domain (the kinase domain) is located in the C-terminal 20 kDa. What the N-terminal 449 kDa of the protein is required for is still not clear despite decades of dedicated work on this issue. In normal cells, DNA-PK<sub>cs</sub>, unlike ATM, is present as a monomer. DNA-PK<sub>cs</sub> is recruited to a DSB through its interaction with the C-terminal end of a DNA-bound Ku86 (131, 132). DNA-PK<sub>cs</sub> recruitment is followed by the inward translocation of Ku complex, which promotes homodimerization of the DNA-PK<sub>cs</sub> molecules bound to separate ends of a DSB. The

end result is that the two ends of the DSB are held together by the DNA-PK<sub>cs</sub> homodimer (168). The binding of each DNA-PK<sub>cs</sub> monomer to Ku and the DNA also results in DNA-PK<sub>cs</sub> acquiring a higher affinity for the DNA, which, in turn, stimulates the kinase activity of DNA-PK<sub>cs</sub> (169-171). The initial phosphorylation target of DNA-PK<sub>cs</sub> appears to be itself, and this autophosphorylation is critical for productive C-NHEJ (172, 173). The autophosphorylation of DNA-PK<sub>cs</sub> happens both *in trans and cis*, but the biological significance of this distinction is unknown (174). Regardless, autophosphorylation seems to induce a subsequent conformational change that permits downstream processing steps and simultaneously also induces dissociation of DNA-PK<sub>cs</sub> (175, 176). The dissociation of DNA-PK<sub>cs</sub> is a mechanistically important step, because it allows the subsequent downstream processing events to proceed. Mutants of DNA-PK<sub>cs</sub> that cannot be autophosphorylated, remain bound to DSB ends and actually result in a more severe IR-sensitive phenotype than the null mutants (177). As importantly, the activation of DNA-PK<sub>cs</sub> also results in the phosphorylation (and activation) of downstream targets. As an added complexity, however, it is still unclear if DNA-PK<sub>cs</sub> is directly or indirectly required for these phosphorylation events. For example, Ku, XRCC4, XLF, and LIGIV are all phosphorylated robustly *in vitro* by DNA-PK<sub>cs</sub>, but phosphorylation of these proteins by DNA-PK<sub>cs</sub> seems to be non-essential *in vivo* and/or carried out by a different kinase (125, 176, 178-180). The best example of this conundrum is Artemis. Artemis is intensely phosphorylated by DNA-PK<sub>cs</sub> *in vitro* (181, 182); in contrast the *in vivo* phosphorylation of Artemis seems more dependent on ATM (183-185).

Artemis was first identified as the gene defective in rare subgroup of RS-SCID (radiosensitive severe combined immunodeficient) patients. Artemis is a single-strand-specific 5' to 3' exonuclease. Importantly, after phosphorylation by ATM or DNA-PK<sub>cs</sub>, Artemis acquires an endonuclease activity specific for hairpin and overhang structures. This enzymatic activity is required for the hairpin opening and overhang processing of the coding ends generated by RAG1/RAG2 recombinase during V(D)J recombination (182). Mechanistically, this explains why Artemis-deficient patients present with a profound immune deficiency — they are incapable of completing V(D)J recombination. The radiation sensitivity of Artemis patients is much less well understood, but is

presumably due to certain types of DNA ends that are generated by IR exposure, which can't be religated without some sort of (Artemis-dependent) overhang processing event (186).

In summary, DNA-PK<sub>cs</sub> and Artemis are critical components of the C-NHEJ pathway and together they generate a nuclease activity that is critical for V(D)J recombination and DNA DSB repair.

### **LIGIV:XRCC4:XLF**

Ku, as noted above, is required for the recruitment and activation of DNA-PK<sub>cs</sub> to DNA DSBs. Ku is also required for the recruitment of the C-NHEJ ligation complex (which is composed of LIGIV and its two accessory factors, XRCC4 and XLF), because the ligation complex by itself does not have particularly high affinity for DNA DSBs (135, 136). The ligation complex is formed in a stoichiometry of 2:2:1 (XRCC4:XLF:LIGIV) and the complex is well conserved throughout evolution (187, 188). Budding yeast, for example, have orthologs of all three components, named Lif1:Nej1:Dnl4, respectively (15, 189).

LIGIV is one of three mammalian ligases (190). Interestingly, the other two ligases, LIGI and LIGIII, do not appear to operate at all in C-NHEJ. LIGI appears to have the most wide-spread usage with documented roles in DNA single-stranded break repair (191), Okazaki fragment processing during DNA replication (192) and HR (193). LIGIII, in contrast, is used exclusively in A-NHEJ ((194); see below). Importantly, neither ligase appears capable of substituting for LIGIV during C-NHEJ (158). LIGIV is a 911-amino acid protein and was first identified through a homology search using a C-terminal conserved region of the known eukaryotic ligases at the time, LIGI and LIGIII (195). A unique feature of LIGIV is a large C-terminal region containing two BRCT domains. Interestingly, LIGIV interacts with the XRCC4 homodimer not through one of the BRCT (which are known protein:protein interaction modules) domains but through a inter-BRCT linker domain (196). Structural studies have demonstrated that one molecule of LIGIV is asymmetrically bound to the same helical region of two polypeptides within the XRCC4 dimer (188). The second BRCT domain, although dispensable for direct binding, is nonetheless important to stabilize LIGIV:XRCC4 complex (196). The first

BRCT domain interacts with Ku (197), and this interaction, as noted above, increases the affinity of the ligation complex to DNA (135, 136).

XRCC4 was first identified as the human cDNA that rescued the defects of the XR-1 CHO (Chinese hamster ovary) cell line, which is defective for DSB repair and V(D)J recombination (198). XRCC4 is a 334-amino acid protein with a globular N-terminal head domain followed by a long helical tail. XRCC4 stabilizes LIGIV (199), and functions as a scaffold protein by recruiting LIGIV and XLF to the break site. Consequently, XRCC4-null mutations phenocopy LIGIV-null mutations and this mechanistically explains the attendant IR sensitivity and V(D)J recombination defects. XRCC4 can form a filamentous structure through head-to-head interactions *in vitro*, but the *in vivo* importance of this aggregation activity is unknown. Interestingly, however, the interaction of LIGIV with the XRCC4 inter-BRCT linker domain weakens the XRCC4 filamentous structure and dissociates it into a dimer form (187, 200). Finally, XRCC4 also shows a strong interaction with DNA-PK<sub>cs</sub>, albeit not with Ku (138, 201, 202). This interaction presumably also helps to anchor the ligation complex at the DSB site.

Two different groups identified XLF (also called Cernunnos). One group complemented cells from RS-SCID patients with a cDNA library and found that a XLF cDNA could rescue the IR sensitivity and V(D)J recombination defects of a XLF patient-derived cell line (203). The other group isolated XLF from yeast-two hybrid screens as a protein that interacted with XRCC4 (204). XLF is a 299-amino acid protein and it has a high structural similarity with XRCC4 despite a low (13.7%) sequence identity (204). XLF consists of a globular head domain and an elongated coiled-coil region similar to XRCC4, but its coiled-coil tail is shorter than that of XRCC4. XLF is not essential for C-NHEJ but it greatly stimulates the ligation of non-compatible ends by promoting re-adenylation of LIGIV (205, 206).

Unlike the other C-NHEJ factors, the ligation complex components are essential in mice. It seems likely that LIGIV is also essential for humans, because all the described LIGIV patients have hypomorphic mutations (see below). Interestingly, however, LIGIV is not essential in either mouse (207) or human somatic cell lines (207), suggesting a developmental role for the gene. Indeed, LIGIV-null mice are embryonic lethal due to



the p53-mediated neuronal apoptosis (208-210). Actively dividing neuronal cells apparently generate higher levels of ROS than most cells, and they seem to be unable to manage ROS-induced DNA breaks without functional LIGIV. With a phenotype and timing very similar to LIGIV, XRCC4-null mice are also late embryonic lethal (211). Confusingly, XLF-null mice are viable (212) and this has led to suggestions that perhaps only LIGIV and XRCC4 constitute the ligation complex, whereas XLF performs some non-essential, ancillary role. This issue is under active investigation from a number of laboratories.

### **The polymerase X family**

There are six polymerase X family members and three of them are required for C-NHEJ; pol  $\mu$ , pol  $\lambda$ , and TdT (terminal deoxynucleotide transferase) (134). All three enzymes share the polymerase X core catalytic homology domain as well as including an enzymatic activity that is capable of adding dNTPs to the 3'-end of DNA. Pol  $\mu$  and pol  $\lambda$  are capable of both DNA-template-dependent and DNA-template-independent modes of synthesis (213). Impressively, they can even synthesize DNA across a gap (214). In C-NHEJ, the lack of homology at the DNA ends to be joined probably generates gaps at a high frequency, which probably explains the requirement for polymerases like pol  $\mu$  and  $\lambda$ . TdT is a lymphoid-specific gene expressed only in pre-B and pre-T cells and it synthesizes DNA on the 3'-hydroxyl of a short overhang in a template-independent manner, which makes it useful for increasing the functional variability of V(D)J recombination products (215).

### **Mechanism of C-NHEJ**

Although the exact mechanism of C-NHEJ is not fully elucidated, a reasonable model of the repair mechanism exists:

#### 1) Ku70/Ku86 binding

The Ku70/Ku86 complex is first recruited to DNA DSB ends and serves as a damage sensor. Ku70/86 is an abundant protein in cells and it has a very high affinity for DNA ends, which enables free DNA DSB ends to recruit Ku within 10 sec of a DSB forming

(216, 217). The Ku heterodimer, which forms a ring-shaped structure, binds to dsDNA by encircling it such that the DNA passes through Ku like a thread through a bead (129).

#### 2) DNA-PK<sub>cs</sub> binding

DNA-PK<sub>cs</sub> is recruited to a DSB through the interaction with the C-terminal end of DNA-bound Ku86 (131). The DNA-bound Ku complex then translocates inward and allows DNA-PK<sub>cs</sub> to bind at the very end of the DSB (170). DNA-PK<sub>cs</sub> makes direct contacts with DNA termini, occupying about a 10 nt region proximal to the free end, and this interaction with DNA stimulates a conformational change within DNA-PK<sub>cs</sub>, which subsequently activates the kinase activity of DNA-PK<sub>cs</sub> (169, 170, 218). Importantly, DNA-PK<sub>cs</sub> forms a homodimeric synaptic complex with the other DNA-PK<sub>cs</sub> molecule bound to the opposite side of the break. This interaction stimulates *trans*-autophosphorylation across the DSB and is also presumed to function in end bridging.

#### 3) DNA end processing

The structure of DSB ends varies depending on many parameters and end processing is often required to generate ligatable ends. For example, IR-induced DNA breaks can have blunt or staggered ends with 5'- or 3'-overhangs, and these ends may need to be processed by either the MRN complex, Artemis or both. In V(D)J recombination, the RAG1/RAG2 recombinase makes DSBs with hairpinned ends and the processing of these ends requires Artemis and (directly or indirectly) DNA-PK<sub>cs</sub>. Finally, in order to ligate DNA ends together, the LIGIV complex needs a 3'-phosphate group. If the 3'-phosphate group is lost through DNA damage or end processing, polynucleotide kinase (PNK) phosphorylates the 3'-end and enables ligation (219, 220). This latter event apparently happens infrequently, since the loss-of-function of PNK generates, at best, a mild C-NHEJ defective phenotype (221).

#### 4) DNA end ligation

The LIGIV:XRCC4:XLF complex is recruited to DSBs using interactions with either Ku or DNA-PK<sub>cs</sub> (135, 136, 197, 222). Once the ligation complex is recruited to a DSB end, it induces an inward translocation of Ku, freeing the DNA end for ligation (223).

The active site lysine of *LIGIV* needs to be adenylated for ligation, and most of *LIGIV* in cells is present as an adenylated form. This adenylate group, adenosine 5'-monophosphate (AMP), is transferred to the 5'-phosphate of the DNA substrate during ligation. Adenylation of the DNA activates the 5'-phosphate for the nucleophilic attack by the 3'-hydroxyl group, which displaces the AMP and covalently joins the ends together (190). The role of *XRCC4* is to stabilize *LIGIV* and to help with the adenylation of *LIGIV*. Lastly, *XLIF* facilitates the re-adenylation of *LIGIV* and it promotes the ligation of non-compatible ends (205, 206, 224).

Even though C-NHEJ has been extensively studied for the past two decades, there are still a few critical questions that have yet to be adequately answered. First, how C-NHEJ components are dissociated from a repaired DSB is perhaps the least understood topic. In particular, the Ku heterodimer is a donut-shaped protein that encircles the DNA. Because DNA-PK<sub>cs</sub> and the ligation complex are recruited by Ku (223), it is highly unlikely that Ku is released before ligation. This however, presents a large logistical problem, because the Ku subunits are topologically interdigitated (129) and cannot just disassemble like many other multi-subunit complexes. Thus, following successful C-NHEJ-mediated DNA DSB repair the Ku heterodimer is presumably topologically trapped on the chromosome. If this Ku is not removed, it would almost certainly subsequently interfere with dynamic chromosomal processes such as DNA replication and RNA transcription. One possible way to solve this problem would be to proteolytically degrade the Ku ring off of the DNA (225) — after DSB repair has been completed — in a manner analogous to the disassembly of the cohesion ring from sister chromatids following DNA replication. Unfortunately, there is no clear evidence for the involvement of proteases in Ku dissociation yet.

Another area in the C-NHEJ mechanism that still requires clarification is what factors are required to activate the repair process. Thus, C-NHEJ is usually described (as I have done) as a sequential event, but it is not clear whether all the components need to be recruited first to initiate C-NHEJ or not. For example, in yeast, *Dnl4* and *Lif1* are required to initiate NHEJ (226, 227). Similarly, not only Ku and DNA-PK<sub>cs</sub>, but also

LIGIV and XRCC4 can suppress other repair pathways (228), implying that LIGIV and XRCC4 might be involved in very early step of C-NHEJ.

### **Human diseases associated with C-NHEJ defects**

Work on C-NHEJ has flourished over the past two decades and it continues to be a vibrant and highly competitive area of research. One of the reasons for this interest has been the realization that defects in C-NHEJ genes are directly responsible for a veritable bevy of serious human diseases. In many cases the complete phenotypes are quite complex — too complex for an exhaustive discussion here — but a brief synopsis will be useful. Ataxia-telangiectasia (A-T) is a rare, autosomal recessive syndrome caused by mutations in ATM and it displays multisystem defects such as telangiectasia (a condition characterized by the dilation of capillaries), cerebellar degeneration, immunodeficiency, genomic instability and cancer predisposition (229). Mutations in the other major upstream signaling kinase, ATR, cause Seckel syndrome, which is characterized by microcephaly and developmental delay (230).

The subunits of the MRN complex have also been identified as disease-causing alleles. NBS1-mutations, which lead mostly to truncated Nbs1 variants, cause the Nijmegen chromosome breakage syndrome, which is an autosomal recessive chromosomal instability disorder characterized by premature aging, microcephaly, growth retardation, immunodeficiency, cancer predisposition and IR sensitivity (57, 58, 231, 232). The NBS1 gene is located on human chromosome 8q21 and it codes for a protein termed nibrin. Over 90% of NBS1 patients are homozygous for a founder mutation: a deletion of five base pairs, which leads to a frame-shift and C-terminal protein truncation (233). Given that MRN is required for the recruitment of ATM to a DSB end, it is perhaps not at all surprising that hypomorphic mutations in Mre11 cause a disease termed ataxia-telangiectasia-like disorder (ATLD). ATLD patients display ataxia and neurodegeneration, similar to the symptoms of A-T patients (234, 235). A single patient with a hypomorphic mutation in Rad50 has been reported and the patient, again not surprisingly, presented with phenotypes similar to NBS1 patients (59). Thus, mutations to any of the subunits of MRN, ATM or ATR result in profound pathological

disorders in humans, highlighted by neurological and immunological dysfunction and cancer predisposition.

Recently, the first DNA-PK<sub>cs</sub> human mutation was described in a T<sup>-</sup>B<sup>-</sup> RS-SCID patient (236). This patient displays a classic SCID phenotype similar to RAG1-, RAG2- or Artemis-deficient patients. Confusingly, however, the putative hypomorphic missense mutation identified in the DNA-PK<sub>cs</sub> patient affected neither protein expression nor (auto)phosphorylation, but rather appeared to cause a defect in Artemis activation (236). The presumption, therefore, is that this mutation disrupts a DNA-PK<sub>cs</sub> domain required for the DNA-PK<sub>cs</sub>:Artemis interaction. Direct experimental evidence for this hypothesis, however, is lacking. Mutations in Artemis also lead to a RS-SCID syndrome, which is characterized by a hypersensitivity to DNA DSBs and an absence of functional B- and T-lymphocytes (237).

Patients resulting from defects in the C-NHEJ ligation complex components are extremely rare. For example, to date, only seven LIGIV patients have been reported. One of these patients did not have any obvious developmental defects but subsequently developed a radiosensitive leukemia (238). Tragically, this patient had an extreme (and fatal) reaction to the therapeutic radiation treatment. A cell line derived from this patient, 180BR, has a mutation within a highly-conserved motif encompassing the active site of LIGIV, and the resulting mutant protein is severely compromised in forming an enzyme-adenylate complex *in vitro* (238). Four other patients with mutations in LIGIV presented with phenotypes (*e.g.*, chromosomal instability, developmental and growth delay, unusual facial features, pancytopenia and various skin abnormalities) similar to ATLD/NBS1 patients (239). Unlike NBS1-mutant cell lines, however, LIGIV-mutant cell lines show normal cell cycle checkpoint responses (240, 241). These LIGIV syndrome patients appeared to have suffered sporadic mutations as each had a unique homozygous or compound heterozygous hypomorphic mutations, which did not completely abolish — but did significantly reduce — enzyme function (242). A sixth LIGIV patient presented with an acute T-leukemia (240). This radiosensitive leukemia patient was a 4.5-year-old boy with a facial shape (bird-like) strongly reminiscent of NBS1 patients. Importantly, however, this patient had no mutations in their Nbs1 genes, but did have a homozygous truncating mutation in LIGIV (240). The seventh LIGIV patient had a T<sup>-</sup>B<sup>-</sup>NK<sup>+</sup> RS-

SCID phenotype without developmental defects. Somewhat surprisingly, a cell line derived from this patient, SC2, had no detectable LIGIV protein and no detectable DNA end joining activity (243), suggesting that it might be functionally null for LIGIV. More extreme than LIGIV is XRCC4, for which no patients have been described, suggesting that, as in the mouse (211), the gene is essential. In contrast, XLF patients have been described and they universally show growth retardation, microcephaly, and an immunodeficiency characterized by a profound B- and T-cell lymphocytopenia (203). The more frequent occurrence of XLF patients compared to LIGIV and XRCC4 patients is consistent with the hypothesis that LIGIV:XRCC4 is the core ligation complex and XLF only provides accessory activity.

### **Alternative non-homologous end joining (A-NHEJ)**

A-NHEJ was first documented in chicken DT40 cells as a DSB repair pathway that overlapped with neither C-NHEJ nor HR (159). DT40 cells repair IR-induced DSBs in a highly reproducible biphasic manner. The half time of the fast component is 13 min and the half time of the slow component is 4.5 hr. In DT40 cells lacking the expression of Ku, the fast component disappears but a significant proportion of DSB repair can still be performed using the slow repair pathway. Co-inactivation of HR components, such as Rad51B, Rad52, and Rad54 did not change the activity of the slow repair pathway, suggesting that the slow repair pathway did not represent HR (159). Importantly, this slow-kinetic repair pathway was also observed in the 180BR cell line, a primary human fibroblast cell line derived from a patient with an inactivating mutation in LIGIV (238, 244). Thus, the LIGIV deficiency of 180BR compromised the fast, but not the slow, repair pathway in a manner analogous to Ku-deficient DT40 chicken cells. Finally, wortmannin, a DNA-PK<sub>cs</sub> inhibitor, treatment abolished the fast repair pathway in wild-type human cells but not the slow repair pathway. Together, these results implied the presence of a novel repair pathway, which is Ku-, DNA-PK<sub>cs</sub>- and LIGIV-independent (244). The existence of this pathway was subsequently confirmed and it has been termed A-NHEJ.

Following the identification of the A-NHEJ pathway, a large amount of work has gone into establishing what proteins/genes are involved in this process. Using a classical

fractionation:complementation approach, LIGIII was identified as the first authentic A-NHEJ factor (194). By inference, XRCC1 was therefore implicated as well. XRCC1 forms a stable heterodimer with LIGIII in the nucleus and is known to stabilize LIGIII (245-247). Subsequent experimentation confirmed the involvement of XRCC1 and extended them to PARP1 as well. In DNA-PK-defective (*i.e.*, C-NHEJ deficient) rodent cells, DNA DSB repair required the synaptic activity of PARP-1 and the ligation activity of LIGIII:XRCC1 (248). Current models suggest that PARP1, which like Ku has a very high affinity for DNA double-stranded ends (27, 249), competes with Ku for a DNA DSB — if Ku wins, C-NHEJ ensues; if PARP-1 wins, A-NHEJ is utilized. These three proteins (LIGIII, XRCC1 and PARP1) constitute the core A-NHEJ factors, but it is clear that more (perhaps many more) await identification. For example, histone H1 has a stimulatory effect on the A-NHEJ pathway, presumably by facilitating the alignment of the DSB ends (250). In addition, the A-NHEJ pathway is assumed to require at least modest end-resection in a search for microhomology regions and this process is likely dependent on MRN and CtIP (44, 251, 252). Finally, LIGIII interacts with polymerase  $\gamma$  and while this is only known to be required for mitochondrial DNA replication (253) it also suggests the possible role of polymerase  $\gamma$  in A-NHEJ.

Unlike HR and C-NHEJ, LIGIII-dependent A-NHEJ is not evolutionarily conserved and is present only in mammals and *Xenopus laevis* (190). This makes it difficult to use simple model organisms, such as yeast, for A-NHEJ study. In addition, the physiological function of the A-NHEJ pathway is unclear. That is, the loss of HR leads to profound meiotic defects and the loss of C-NHEJ leads to acute IR sensitivity and severe V(D)J recombination defects, but what pathological states the absence of A-NHEJ generates is not known. Thus, there have been several attempts to establish systems where A-NHEJ activity is measurable. Because A-NHEJ is suppressed by C-NHEJ (254), a C-NHEJ deficient condition is frequently used to measure A-NHEJ activity. Moreover, the sequence of the repaired junctions has been used as an indicator of which pathway is being used, because HR, C-NHEJ, and A-NHEJ use micro-homology differently (89, 254, 255). Two assays that have gained popularity are a system in which the expression of a truncated RAG2 generates robust A-NHEJ activity during V(D)J recombination

(251) and a plasmid end-joining assay in which a substrate, designed to detect only microhomology-mediated end joining, is used (252).

The mechanism of A-NHEJ is still unclear, but consideration of the few known A-NHEJ components gives some hints. First, the MRN complex and ATR signaling are required for end-resection (252, 256-259). The end-resection process in A-NHEJ seems very similar to that in HR, with the exception that Rad51 appears not to be involved in A-NHEJ (260). It is unknown how the degree of resection is regulated and how/when the homology pairing that is required for repair happens. The involvement of histone H1 around the damage site may be required to keep the broken ends proximal to one another (250). This is because histone H1 facilitates PARP-1 activity and PARP-1 is thought to be bound to the DSB ends and likely mediates their synapsis. About the only thing known for sure is that the LIGIII:XRCC1 complex clearly covalently links the ends of the DSB back together and somewhat surprisingly this does not appear to be a function that can be assumed by either of the other two ligases. Finally, it is clear that nucleases (perhaps MRN and CtIP) and polymerases (perhaps pol  $\gamma$ ) will be needed, but their identity remains elusive.

### **LIGIII**

Unlike other ligases, LIGIII shows molecular heterogeneity, implying that it may have diverse roles. Alternative splicing generates two different LIGIII isoforms ( $\alpha$  and  $\beta$ ), which are expressed with or without a C-terminal BRCT domain, respectively. The  $\alpha$ -form has a C-terminal BRCT motif and is expressed ubiquitously. The  $\beta$ -form of LIGIII, which does not have the BRCT domain, is expressed exclusively in germ cells. Because LIGIII $\beta$  appears to have a very specialized and constrained role, the rest of this Introduction will be concerned only with LIGIII $\alpha$ , hereafter referred to simply as LIGIII. LIGIII interacts with XRCC1 and this interaction occurs through LIGIII's BRCT domain. Additional heterogeneity comes from alternative translational initiation of LIGIII, which results in mitochondrial- and nuclear-specific forms based on the presence and absence of an N-terminal mitochondrial leader sequence (MLS), respectively (246, 261, 262).

The nuclear LIGIII isoform forms a heterodimeric complex with XRCC1 (245), and this LIGIII:XRCC1 complex also functions in single-strand break repair (SSB), short



patch base excision repair (BER) and nucleotide excision repair (NER) (190) in addition to A-NHEJ. LIGIII and XRCC1 interact through BRCT domains (263) and this interaction seems to result in mutual stabilization (247, 264). XRCC1- and LIGIII-deficient mice are early embryonic lethal at 6.5-7.5 days postcoitum (dpc) and 8.5 dpc, respectively (265, 266). The similarities in the phenotypes of these knockout mice suggest strongly that the two genes always work together. This belief, however, is likely to be true only in the nucleus. Thus, unlike LIGIII, XRCC1 is not present in mitochondria, implying that LIGIII has a XRCC1-independent function in mitochondria (267, 268). How mitochondrial LIGIII remains stable in the absence of XRCC1 is not well understood.

Recently, Maria Jasin's group generated a LIGIII-null mouse model (269). LIGIII-null mice were early embryonic lethal (266), so a conditionally-null LIGIII mouse strain was generated (269). Surprisingly, LIGIII-null MEF cell lines complemented with a mitochondrial-only expressed LIGIII cDNA were alive and neither showed any growth defect nor sensitivity against DNA damaging agents. This observation conclusively demonstrated that the essential function of LIGIII resides in the mitochondria and is presumably related to mitochondrial DNA replication. Perhaps even more surprisingly, the expression of a LIGI or *Chlorella* virus DNA ligase (the latter constitutes the minimal eukaryal nick-sealing enzyme (270)), cDNA in mitochondria was also sufficient to suppress the lethality of LIGIII loss-of-function. The essential role of LIGIII in mitochondria is further supported by the recent discovery that LIGIII is required for mitochondrial Okazaki fragment maturation (271). All of these studies, however, beg the pressing question of what role — if any — LIGIII plays in the nucleus. Recent studies have suggested that one of the main physiological functions of nuclear LIGIII may be in regulating chromosomal translocations (228, 272). This result had actually been anticipated since the sequencing of human cancer genomes demonstrated that many of the chromosomal translocations observed in cancer cells exhibited the preferential use of microhomology (273), which, of course, is the salient feature of A-NHEJ. To measure the contribution of C-NHEJ and A-NHEJ on chromosomal translocation formation, two different fluorescent markers were incorporated into different chromosomal locations with adjacent *I-SceI* sites in LIGIV- or XRCC4-deficient mammalian cells. After *I-SceI*

expression, the cleavage and translocation frequency was scored by identifying derivative chromosomes that contained both markers. These studies demonstrated that LIGIV and XRCC4 were not only NOT required for the translocations, but that they actually normally suppressed the translocations because their deficiency increased the translocation frequency and did not change the sequence characteristics of the translocation junctions (272). These data compellingly suggest that A-NHEJ, and not C-NHEJ, is involved in mediating chromosomal translocations (228). The utility of the conditionally-null LIGIII cell line was evident in another study where specific chromosomal breaks were induced using a zinc-finger nuclease, and — expectedly — the translocation efficiency in cells not expressing LIGIII was dropped (272).

Finally, it should be emphasized that although the field now readily accepts the existence of a Ku-, DNA-PK<sub>cs</sub>- and LIGIV-independent and LIGIII-dependent A-NHEJ pathway, it is very unclear how many microhomology-mediated end-joining (MMEJ) pathways actually exist (274). In addition to the A-NHEJ pathway described above, there are indications of other variations of MMEJ based on the requirements for different protein factors (274, 275).

### **The relationship between DNA DSB repair pathways**

The mechanism of pathway choice is currently perhaps the most actively researched issue in the field of DSB repair. That competition between HR and C-NHEJ for DSB ends can occur is widely accepted. This competition is perhaps best illustrated by studies investigating gene targeting, which is an exclusively HR-mediated process (83, 276-278). In mammalian cells, the gene targeting frequency increases 5- to 20-fold in Ku-depleted cells, indicating that C-NHEJ normally suppresses HR (254, 276). The binding affinity and the abundance of end-binding factors (*e.g.*, Ku or MRN) at DSB ends seems to play an important role in deciding what repair pathway (*e.g.*, C-NHEJ or HR, respectively) is utilized for a particular DSB. However, it is not clear whether a single end-binding factor is enough to commit a broken end to a specific pathway or not. For example, in yeast, the Dnl4:Lif1 complex is required to stabilize DNA-bound yKu and thus is also likely to be required for pathway choice in that organism (227). In an analogous fashion, studies

with human cell extracts also showed that LIGIV:XRCC4 may be required for C-NHEJ initiation (226, 279).

The identification of A-NHEJ as a mechanistically and genetically distinct pathway has only lent additional complexity to this problem. Under normal physiological conditions, A-NHEJ is virtually undetectable. In contrast, A-NHEJ processes more than half of the DSBs under C-NHEJ defective conditions (159). This implies that C-NHEJ and A-NHEJ might also compete against each other for DNA ends, but that under normal conditions, C-NHEJ almost always wins out. Together, these studies suggest that C-NHEJ may be able to suppress both HR and A-NHEJ, whereas HR and A-NHEJ generally do not influence each other (159).

This latter point is likely related to the fact that both HR and A-NHEJ require end resection as a part of their repair process. It seems plausible (although certainly not proven) that the main distinction between the two pathways (*i.e.*, in deciding which pathway to use to repair a DSB) is probably directly related to the degree of resection that occurs, with longer resection favoring HR over A-NHEJ and shorter resection favoring A-NHEJ over C-NHEJ (280-283). Again, how or what limits end resection is not known. The cell cycle seems to be one of the possible regulators. HR is used predominantly during late S and G2 phases when homologous sequences from a sister chromatid become available (283, 284). C-NHEJ is active in all phases of the cell cycle (21), but it plays its most important role generally only during the G1/G0 phases. In contrast, A-NHEJ is also active in all phases of the cell cycle, but it is most active in G2 phase and its activity is compromised when cells enter the plateau phase of growth (285-287). Thus, one model for how pathway choice gets determined would be if the protein expression level and/or post-translational modification of some key DSB repair protein(s) is modified as a function of the cell cycle. Recent studies concerning CtIP are consistent with this hypothesis. End-resection by CtIP is involved in both HR and A-NHEJ during S/G2 and G1 phases of cell cycle, respectively. Moreover, CtIP protein levels are low during G1 phase and high during S, G2 and M phases, even though the gene is consistently transcribed throughout the cell cycle (288). This regulation of CtIP at the protein level occurs in a kinase-dependent manner. Importantly, the phosphorylation of CtIP on Ser327, which requires cyclin dependent kinase activity, facilitates the

interaction with BRCA1 and MRN complex, which subsequently increases the end resection activity of CtIP (289-291). Importantly, A-NHEJ needs smaller amounts of resection than HR, so that while cells expressing a CtIP that cannot be phosphorylated at Ser327 are defective in HR, A-NHEJ remains unaffected (284).

In summary, pathway choice is a critical event in the life of a DSB. The biological outcome (*e.g.*, error-free repair, error-prone repair, or chromosomal translocation) for a DSB end is utterly different depending upon whether HR, C-NHEJ or A-NHEJ, respectively, is chosen to repair it and thus the cell presumably regulates this choice very tightly. It is clear that additional experimentation in this area is required and warranted.

### **Telomeres**

Unlike bacteria, eukaryotes have linear chromosomes that cause end-replication and end-protection problems. The former issue was vocalized initially by James Watson, who described it as the “end-replication problem” (292). He noted that the end of a chromosome being replicated by the lagging strand cannot be synthesized completely, so that every replication cycle, a chromosome should shorten, with an attendant loss of genetic information. This end-replication problem has apparently been solved by the evolution of the telomere, which is a TG-rich repetitive sequence found at the end of each chromosome (293). This TG-repeat is extended by a dedicated reverse transcriptase, telomerase. Telomerase is minimally composed of two proteins; *Tert* (telomerase reverse transcriptase) and *Terc* (telomerase RNA component). *In vivo*, telomerase is likely composed of additional accessory factors (*e.g.*, *Dkcl1*; Dyskeratosis congenita 1, dyskerin; (294, 295)), but not all of them have been identified. In humans, telomerase is constitutively active in germ cells and some cancer cells, but almost somatic cells have undetectable levels of telomerase activity, which leads to a progressive telomere shortening (165, 296). Biologically in humans, therefore, the telomere acts as a buffering region for chromosome end loss and essentially allows for only a limited number of cell divisions. Once a telomere is shortened to a critical point it loses its end protection activity (elaborated below) and the cells stop replicating and enter into senescence. This replicative senescence is accelerated without telomerase (296). Biologically, senescence is important because it provides a huge barrier to replicative immortality. This is a

barrier that cells must overcome in order to become cancerous (297). The simplest way to overcome the telomere shortening that triggers senescence is for a cell to (re)express telomerase. Not surprisingly therefore, about 85% of all human cancer cells are telomerase positive, making this a popular point for potential clinical therapeutic intervention (165, 298). The approximately 15% of cancer cells that are telomerase negative, must still deal with telomere shortening and they do so through a process named alternative lengthening of telomere (ALT), which uses HR to extend telomeres (299, 300). In summary, all linear chromosomes face the problem of how to stably maintain their ends and telomeres are the universal evolutionary solution.

The requirement for end-protection arises because linear chromosome ends are structurally similar to DNA DSBs. For example, a human telomere is composed of (TTAGGG)<sub>n</sub>, repeats ranging in length from 2 to 14 kb. At the extreme end, the 3' G-rich strand overhangs the double-stranded portion by 100 to 300 nt (301). The transition between the double-stranded portion of the telomere and the G-rich overhang bears structural similarities to DSBs. Consequently, unprotected telomeres can provoke cell cycle arrest through ATM or ATR activation, chromosome fusion through C-NHEJ, or chromosome loss using HR (302-305).

Cells use 2 mechanisms to obviate these deleterious outcomes. First, the telomere forms a specialized structure: the 3' G-overhang invades and hybridizes to a duplex part of telomere, making a classic displacement (D)-loop (10, 306). Because this looping is telomere-specific, it is referred to as a t-loop (307). Second, cells have evolved a group of six proteins (termed “shelterin”) that bind specifically to the t-loop structure, forming a proteinaceous cap on the t-loop. The formation of the unusual t-loop and the coating of this structure by shelterin functionally hide the chromosome end from all DNA damage sensors and DSB repair pathways.

The shelterin complex is composed of six factors; TRF1 and TRF2, POT1 (protection of telomeres 1), TPP1 (tripeptidyl peptidase 1), TIN2 (TRF1 interacting nuclear factor 2), and Rap1 (repressor/activator protein 1) (107). TRF1 and TRF2 bind to the double-stranded telomeric DNA and POT1 binds to the single-stranded DNA. These proteins are held together by the linker proteins, TPP1 and TIN2. TPP1, in particular, ensures that POT1 is connected to the rest of shelterin and facilitates the ssDNA-binding activity of

POT1, which protects the telomere from ATR activation (308, 309). Rap1 is bound to TRF2, but the telomeric defects in mouse Rap1 knockout animals are actually quite mild (310). The specific functions of each of the shelterin component have been principally studied using mouse knockout mutational analyses. For example, a deficiency in TRF2 removes most of the proteinaceous cap from the telomeres and results in ATM activation and subsequent cell cycle arrest through up-regulation of p53 (311, 312). Analogously, POT1 competes with RPA for ssDNA binding to the telomeric G-overhang (313). When POT1 is absent, RPA is recruited to the telomere and initiates an ATR-mediated DNA damage signaling pathway (314). In summary, the end protection problem revolves around a specific structure (t-loops) and a specific complex of proteins (shelterin) that evolved to essentially hid chromosome ends from the DSB repair pathways (307).

In addition to the shelterin complex, there are other proteins associated with telomeres. Paradoxically — at least superficially — these consist mostly of DNA repair and damage signaling proteins. For example, Ku is recruited to telomeres through its interaction with TRF1 and TRF2 in human (112, 139, 141). Superficially, one might think this would facilitate telomere fusions via C-NHEJ, but it appears that at telomeres, Ku's major function is to suppress the aberrant action of HR on the t-loop (117) and thus serves a protective role. DNA-PK<sub>cs</sub>, which can also be recruited to telomeres, may facilitate this aspect of Ku's protective telomeric role (119). The recruitment of the MRN complex to telomeres in higher eukaryotes was first found in meiotic human fibroblasts (48) and later demonstrated to occur via its association with TRF2 (45). Cells from MRN (*e.g.*, NBS1) patients have short telomeres (46, 315). Like Ku, MRN is now realized to play a positive role in telomere maintenance. In particular, DNA end resection is known to be required for the formation of the G-overhang, which, in turn, is critical for t-loop formation (47). Additional experimentation has demonstrated that the nuclease activity of the MRN complex is needed to generate long G-overhangs and thus the MRN complex is required for telomere maintenance (46).

Despite all the protective systems in place to correctly sequester telomeres and to avoid telomere uncapping, occasionally bad things can happen to good telomeres. When such untoward events transpire, the deprotected telomere is recognized as a DSB and it initiates canonical checkpoint signaling and becomes a substrate for DSB repair (316-

318). For example, in the absence of TRF2, the MRN complex processes telomeric ends to such an extent that it activates the ATM:Chk2 signaling pathway (311, 319, 320). Subsequently, these unprotected telomeres form TIFs (telomere-dysfunction induced foci), which include virtually the same factors found in IRIFs (IR-induced foci) such as p53,  $\gamma$ H2AX, MDC1, and 53BP1 (320). Ultimately the TIFs activate the C-NHEJ pathway, which results in LIGIV-dependent telomere fusions. Perhaps the strongest argument for the importance of proper telomere maintenance is provided by the disease Dyskeratosis congenita. Patients afflicted with this disorder have profound bone marrow failure, cancer predisposition and usually succumb before their third decade of life (295). Seven genes (when mutated) have been identified as causing dyskeratosis congenita and all of them are involved in some form of telomerase biogenesis (*e.g.*, hTERT) or telomere maintenance (*e.g.*, TINF2) (294, 295, 321). These studies unequivocally underscore the importance of proper telomere maintenance.

### **Recombinant adeno-associated virus (rAAV)**

As noted above, significant amounts of telomere research have been carried out in simple model systems such as yeast or in genetically modified mice. In neither of these systems, however, is the biology of human telomere maintenance accurately represented (318, 322, 323). Consequently, our laboratory has tried to develop molecular technologies that could be applied to human somatic cells in an effort to generate a potentially more applicable model system. The most important of these technologies is the use of rAAV for gene targeting purposes.

A reliable exogenous gene delivery system is crucial for clinical experiments and for use as a therapeutic tool. Plasmids and retroviruses have been widely used for gene delivery, but there is a limit to these vectors as they have the serious drawback that they integrate almost randomly into the host cells' genome (324). This high random integration rate results in a variety of problems. First, the expression level of an exogenous gene cassette can vary depending on the locus where it is integrated (325). Second, random integrations can cause transformation by generating deleterious mutations (326). Third, the integrated gene-expression cassette tends to be unstable and can be lost over time (327). Currently, the only way to solve all these problems related to

the random integration of gene expression vectors is use the alternative strategy of gene targeting, which has none of these attendant disadvantages. Unfortunately, however, the gene targeting frequency in human cells is generally only on the order of one in  $10^6$  or lower, principally because C-NHEJ (which mediates the random integration) dominates over HR (which mediates the gene targeting events) (278, 328). This all changed, however, in 1998 when Russell and colleague reported the surprising observation of one in  $10^2$  gene targeting efficiency in human cell lines using a rAAV (recombinant AAV) vector (329). Since then, dozens of investigators, our laboratory foremost among them, have validated and extended the use of rAAV. As of today, over 90 different genes have been genetically modified at their respective endogenous loci via rAAV gene targeting methods (330). Thus, rAAV has proven itself over the past decade to be a robust tool with which to modify the genomes of human cells.

Adeno-associated virus type 2 (AAV2) is a human *parovirus* with a 4.7 kb ssDNA genome. AAV2 encodes two ORFs, *rep* and *cap*, which encode replication and capsid proteins, respectively (331). Wild-type AAV preferentially integrates into the human genome at a specific region on chromosome 19, named AAVS1, and stays latent at this location. Products from a helper virus, such as *adenovirus* or *herpes simplex virus*, are required for a productive replication, which allows AAV to enter into its lytic cycle (332). The site-specificity of its integration and the fact that AAV has no known pathological effects made AAV one of the best early vector candidates for gene therapy in the clinic (333). To use AAV as a gene targeting vector, the *rep* and *cap* ORFs are replaced with human genomic DNA corresponding to the gene of interest. When *rep* and *cap* are subsequently provided in *trans* with some additional helper constructs, active virus can be produced that can be used to infect virtually all somatic human cells (334).

rAAV vectors can be used to make substitutions, insertions, and deletion mutations as well as to epitope-tag endogenous loci (335-337).



## **CHAPTER II**

**LIGIV is not essential for the survival  
of human somatic cells**

C-NHEJ (classic non-homologous end joining) is the predominant DNA DSB (double-strand break) repair pathway in human cells. Seven genes, Ku70, Ku86, DNA-PK<sub>cs</sub>, Artemis, LIGIV, XRCC4 and XLF have been identified as key C-NHEJ components. While these proteins are needed for C-NHEJ activity, several of them have additional functions. One prominent example is the telomere maintenance mediated by Ku70:Ku86, which is essential for cell survival. In contrast to Ku70:Ku86, however, another C-NHEJ factor, LIGIV, has no known function in telomere maintenance and appears to function exclusively in C-NHEJ. Importantly, a viable LIGIV-deficient human B-cell line (NALM-6) has been described. Together these observations strongly suggest that C-NHEJ and LIGIV are dispensable whereas telomere maintenance and Ku70:Ku86 are essential for human somatic cell survival. This hypothesis, however, has its detractors. For example, LIGIV patients have been reported but none of them are LIGIV-null, which has been interpreted to mean that LIGIV is essential (at least during development). Moreover, lymphocyte cell lines like NALM-6 are known to depend more on HR (homologous recombination) than C-NHEJ for DSB repair and this has been interpreted to mean that the viable phenotype of the LIGIV-null NALM-6 cells might be a cell line specific phenomenon and not generally applicable to most human cells. To address some of these issues, we inactivated — using rAAV (recombinant adeno-associated virus) -mediated gene targeting — LIGIV in a fibroblastic human somatic cell line (HCT116, a human colorectal cancer cell line) to investigate whether C-NHEJ is an essential process or not. Our LIGIV-null cell line was also viable, confirming that the gene and C-NHEJ are not essential. However, we did observe a strong disequilibrium during the second round of gene targeting. In addition, the cell line displayed a mild growth retardation, increased genomic instability, severe defects in DNA end-joining and an extreme sensitivity to DNA damaging agents. All of these observations suggest that LIGIV (and by extension, C-NHEJ) is not essential for human somatic cell viability but that it does have important roles in maintaining normal cell homeostasis.

## Introduction

A DNA DSB is one of the most deleterious lesions that can occur in cells, because even a single unrepaired DNA DSB can stop the cell cycle and induce cell death (9, 338). To protect themselves from DSBs, cells have developed at least two major DNA DSB repair pathways; HR and C-NHEJ (339, 340). To enact repair, HR uses extensive sequence homology, — generally sequences longer than 30 nucleotides (nt) — and generates repaired products that are generally “error free”. In contrast, C-NHEJ requires only 0 to 4 nt of homology and is consequently more “error-prone”. Recently, a third, minor pathway, A-NHEJ (alternative-NHEJ) that has features reminiscent of both HR and C-NHEJ has been described (286, 341, 342). A-NHEJ, like HR, requires homologous sequences to mediate the repair reaction. In the case of A-NHEJ, however, only 5 to 25 nt of homology (often referred to as “microhomology”) is needed. Additionally, because its reaction mechanism always results in accompanying deletions, A-NHEJ is therefore also like C-NHEJ in that it is “error prone” (274). Depending on the organism and various parameters (*e.g.*, position in the cell cycle), HR, C-NHEJ and A-NHEJ are differentially utilized (21, 254, 274, 340, 343, 344). Lower eukaryotes like yeast utilize HR almost exclusively for all DSB repair events regardless of the cell cycle state (15, 130, 345). In contrast, higher eukaryotes such as humans use C-NHEJ more often than HR and C-NHEJ is the predominant, if not exclusive, repair mechanism utilized during G<sub>0</sub>/G<sub>1</sub> phases (342). This usage bias is, however, not exclusive. For example, during late S and G<sub>2</sub> phases in human cells, HR becomes more active because a proximal homology donor becomes available in the form of a sister chromatid (328). In summary, higher eukaryotes have multiple options available to them in terms of the pathways that can be used to repair a DSB. This pathway choice flexibility is probably beneficial in certain circumstances and has evolutionarily been selected for. It is clear, however, that since each pathway makes a biologically and functionally distinct product, this choice must be very tightly regulated such that the correct product is generated in the correct biological context.

In addition to DSB repair, the C-NHEJ pathway is also required for V(D)J recombination, CSR (class switch recombination), and telomere maintenance (119, 346). V(D)J recombination is the initial step of antigen receptor maturation that occurs in early

B- and T-lymphocytes. Lymphocyte-specific endonucleases, RAG1 (recombination activating gene 1) and RAG2, make DSBs at specific RRSs (recombination signal sequences) and generate biochemically and structurally distinct coding or signal ends, which are subsequently processed and joined via C-NHEJ (2). CSR is an exclusively deletional recombination reaction that replaces one antibody constant region with another, altering the antibody effector function (4). CSR is initiated by a novel recombinase, AID (activation-induced cytidine deaminase), which initially introduces uracil:guanine mismatches into the DNA. These mismatches are recognized either by mismatch repair proteins or UNG (uracil-DNA glycosylase) and eventually converted into DSBs that can be repaired by either C-NHEJ (104) or A-NHEJ (347-349). Telomeres are the ends of linear chromosomes. They consist of a repetitive tract of DNA that assembles into a structure called a t-loop, which is a variation of a classical D-loop (350). The t-loop is also coated by a bevy of proteins that form a proteinaceous cap on the DNA and essentially keep the chromosome end invisible to all of the DNA DSB pathways. The core six proteins that bind in a sequence-specific manner to telomeres are collectively called “shelterin” (107, 108). There are, additionally, a bevy of shelterin-associated proteins found at telomeres and unexpectedly some of these correspond to C-NHEJ factors. For example, the Ku heterodimer, and possibly DNA-PK<sub>cs</sub>, are part of the t-loop structure (120). Interestingly, telomere protection by Ku complex is essential in human cells because Ku loss-of-function mutations provoke cell death triggered by telomere dysfunctions (117, 124). All of the C-NHEJ factors, however, are unlikely to be involved in telomere maintenance and there is, for example, little evidence for a role in LIGIV, or its accessory factors, in telomere maintenance.

In mammalian cells, there are seven identified and well-researched C-NHEJ components: Ku70, Ku86, DNA-PK<sub>cs</sub> (DNA-dependent protein kinase catalytic subunit), Artemis, LIGIV, XRCC4 (X-ray cross complementing group 4), and XLF (XRCC4-like factor; also occasionally referred to as Cernunnos). Ku70 and Ku86 (so named for their approximate molecular weights) form a ring-shaped, highly-interdigitated heterodimer that binds DSB ends in a sequence-independent manner with extremely high affinity (129). The DNA end-bound Ku heterodimer subsequently recruits DNA-PK<sub>cs</sub> using a protein interaction domain found in the C-terminus of Ku86 (170). DNA binding induces

a large conformational change in DNA-PK<sub>cs</sub> and stimulates its kinase activity (169). Activated DNA-PK<sub>cs</sub> subsequently phosphorylates itself (where phosphorylation is likely inhibitory) and a number of other downstream effectors (where phosphorylation is likely activating) (351). One of DNA-PK<sub>cs</sub>'s substrates is Artemis, a structure-specific nuclease involved in end processing (181, 182, 352). Additional DNA-PK<sub>cs</sub> targets may also include LIGIV and/or its two accessory factors, XRCC4 and XLF (189, 353, 354). XRCC4 functions as a scaffold protein for the ligation complex and it stabilizes LIGIV (199). XLF stimulates the ligation of non-compatible DSB ends by facilitating the re-adenylation of LIGIV (205, 206, 355).

Mutation of any of the C-NHEJ components causes many problems. For example, Ku70, Ku86 and DNA-PK<sub>cs</sub> knock-out mice are viable but they present with severe growth defects, severe-combined immune deficiency and profound hypersensitivity to IR (ionizing radiation) (160-162). In humans, the phenotypes are actually more serious. There is no known Ku70 or Ku86 patients, a fact that has been correlated with Ku's essential role in telomere maintenance (115, 117, 124, 276) and only one hypomorphic DNA-PK<sub>cs</sub>-deficient patient has been reported (236). The latter observation suggests that, like Ku, DNA-PK<sub>cs</sub> loss-of-function mutations may not be tolerated in humans. In contrast, Artemis mutant patients with ostensibly null mutations have been described and they lack mature B and T lymphocytes, experience radiosensitive severe combined immune deficiency (RS-SCID) and are predisposed for leukemia (237). LIGIV and XRCC4 knock-out mice are embryonic lethal and representative MEFs derived from these animals show radiation-sensitivity and defects in V(D)J recombination (356). Seven LIGIV patients have been reported so far in the literature and four of them presented with a "LIGIV syndrome", which is characterized by severe radiation sensitivity, chromosomal instability, unusual facial features, developmental and growth delay (239). The other LIGIV patients displayed only a subset of these features but were also afflicted with leukemia and RS-SCID (238). Importantly, however, all seven LIGIV patients described to date had non-null, hypomorphic mutations (230, 239), suggesting that the gene, like Ku70, Ku86 and DNA-PK<sub>cs</sub> is essential.

Michael Lieber's group has studied the function of LIGIV for V(D)J recombination and IR sensitivity in the human B-lymphoid precursor cell line, NALM-6 (207). These

researchers generated, by classical gene targeting technology, a *LIGIV*-deficient NALM-6 cell line, which was — surprisingly — viable. The viability of the cell line was unexpected based upon the reports for *LIGIV* patients described above. Moreover, it was also unexpected based on mice knockout experiments, since *LIGIV*-null animals are embryonic lethal (209, 210, 356). One possible explanation for these results is that the loss-of-function of *LIGIV* may be tolerated specifically only in NALM-6 cells, but is not generally applicable. This hypothesis takes into account the fact that lymphoid cells tend to use more HR than C-NHEJ (357), so the absence of *LIGIV* might be less toxic to NALM-6 than most of the other cell types. One way to address this issue would be by inactivating *LIGIV* in a non-B human somatic cell line.

Accordingly, we report here the inactivation of *LIGIV* in the fibroblastic human colorectal cancer cell line, HCT116, using rAAV (recombinant adeno-associated virus) gene targeting methodologies (334). Two rounds of gene targeting were used to generate a *LIGIV*-null HCT116 cell line, which was viable. The ability to obtain a viable *LIGIV*-null clone is consistent with existence of the *LIGIV*-null NALM-6 cell line and strongly suggests that neither *LIGIV* nor C-NHEJ is essential in somatic cells. Importantly, however, the *LIGIV*-null HCT116 cells were not aphenotypic. First, there was a large asymmetry in the second round gene targeting frequency with 15 out of 16 second-round targeting events occurring on the already inactive allele, indicating that there is a significant disadvantage to being *LIGIV*-deficient. In addition, the *LIGIV*-null HCT116 cells were extremely sensitive to DNA damaging agents and had defective DNA end-joining activity. So, We conclude that *LIGIV* and C-NHEJ are not essential for human somatic cell survival, but they are nonetheless critical for maintaining normal cell homeostasis.

## Materials and Methods

### Cell culture

HCT116 cells (ATCC) were grown in McCoy's 5A media (Mediatech) supplemented with 10% FBS (fetal calf serum; Cambrex) and 100 U/ml penicillin and 100 U/ml streptomycin (Invitrogen). All cells were grown at 37°C in a humidified incubator with 5% CO<sub>2</sub>.

### Targeting vector construction

The targeting vector was constructed as described (334). The right and left homology arms were amplified by PCR using genomic DNA purified from wild-type HCT116 cells as the template. The primers used to make left homology arm of LIGIV targeting vector were P2.1 and P2.2 (see below). The right homology arm was constructed by PCR using the primers P2.3 and P2.4 (see below). As fusion PCR templates, the two arm constructs and a 4 kb *PvuI*-digested fragment of the pNeDaKO-Neo vector (334) were used. Fusion PCR products were amplified by P2.1 and P2.4 primers and then purified by gel-extraction kit (Qiagen). The fusion construct was subsequently *NotI*-digested and then cloned into a *NotI*-digested pAAV-MCS vector.

P2.1

(5'-ATACATACGCGGCCGCGCAGAAACATGCAGTATTTTCCCCTA-3')

P2.2

(5'-GCTCCAGCTTTTGTTCCTTTAGCAAAGCGGTGATGAATCTTCTCGT-3')

P2.3

(5'-CGCCCTATAGTGAGTCGTATTACAGATGGAAAAGATGCCCTCAAAC-3')

P2.4

(5'-ATACATACGCGGCCGCTTGTGTTTTCTGCACTATTTCTATTC-3')

### Packaging and isolating virus

AAV-293 cells (Stratagene) were grown in DMEM media at 37°C in a humidified incubator with 5% CO<sub>2</sub>. The cells were split into a 10 cm culture dish one day before transfection. pAAV-RC and pAAV-helper plasmid from the AAV Helper-Free system

(Stratagene), and the LIGIV targeting vector (8 µg of each) were co-transfected using Lipofectamine 2000 (Invitrogen) following the manufacturer's protocol. Virus was harvested 2 days after transfection by collecting in 1 ml DMEM media and then performing a freeze and thaw cycle three times with vigorous vortexing in between. The resulting cellular debris was clarified by centrifuging at 13,000 rpm for 2 min, after which the virus-containing supernatant was collected and used for a subsequent infection.

### **rAAV Infection**

HCT116 cells were plated on a 6-well plate one day before infection. When the cells were ~ 60 to 70% confluent, fresh media (1 ml) was added with an adequate amount of virus. After a 2 hr incubation at 37°C, 4 ml of fresh media was added to the virus-containing media. Two days after infection, the cells were trypsinized and transferred at 1000 cells/well on 96-well plates and selection was started using 1 mg/ml G418.

### **Isolation of genomic DNA and genomic PCR**

Genomic DNA for PCR screening was isolated using a Gentra Puregene Cell Kit according to the manufacturer's instructions (Qiagen). The DNA was dissolved in a final volume of 25 µl, 1µl of which was subsequently used in each PCR reaction. For screening, the RArmF/KO3'R primer set and the P1F/P2R primer set (see below) were used for the first and second rounds, respectively.

RArmF (5'-CGCCCTATAGTGAGTCGTATTAC-3')

KO3'R (5'- AAAATGAGACATCATTCACCCCGTGAT-3')

P1F (5'- GGGTTGGAGCAAAACAGTTATTAATGTAG - 3')

P2R (5'- CAATTGAGTCTAAAAGGTCGTTTACTTGC - 3')

### **Immunoblotting**

LIGIV expression was characterized by first preparing nuclear extracts using a CellLytic NuCLEAR Extraction Kit according to the manufacturer's instructions (Sigma). Nuclear extracts (30 µg) were electrophoresed on a 7.5% SDS-polyacrylamide gel (BioRad) and rabbit, anti-human LIGIV antibody (Serotec) was used at a 1:1000 dilution. To screen



for complemented clones, we used whole cell extracts prepared with RIPA buffer (Sigma). LIGIV (Abcam) and HA antibodies (Covance) were used at 1:1000 dilution. As a loading control, an actin antibody (SantaCruz) was used at a 1:250 dilution.

### **Etoposide sensitivity**

Etoposide sensitivity assays were performed as described with a slight modification (358). Three hundreds cells were plated on a 35 mm culture dish ~ 17 to 19 hr prior to drug treatment. Etoposide was dissolved in DMSO (dimethyl sulfoxide) to generate a 10 mM stock solution. The cells were treated with etoposide at varying concentrations and then incubated for an additional 7 to 10 days. The cells were subsequently fixed with a solution of 10% methanol and 10% acetic acid containing crystal violet. A colony was scored as viable when it consisted of 50 or more cells.

### **Microhomology assay**

The microhomology assay was performed as described (23). The pDVG94 plasmid (2.5 µg) was restriction digested with *EcoRV* (NEB) and *AfeI* (NEB), purified and then transfected using Lipofectamine 2000 (Invitrogen). Plasmid DNA was recovered 48 hr after transfection using a modified Qiagen mini-preparation protocol (254). The repaired junctions were PCR amplified using FM30 and 5'-radiolabeled DAR5 primers. The resultant radioactive PCR products were restriction digested with *BstXI* (NEB) and the digested PCR products were separated on a 6% TBE polyacrylamide gel. The gel was subsequently dried and exposed to X-ray film for 10~15 min.

DAR5 (28-mer forward primer: 5'-TGCTTCCGGCTCGTATGTTGGTTGGAAT-3')

FM30 (26-mer reverse primer: 5'-CTCCATTTTAGCTTCCTTAGCTCCTG-3')

### **Complementation**

We constructed a wild-type LIGIV cDNA and LIGIV cDNA with a HA-epitope tag at the C-terminal end. We cloned them into pcDNA3.1(+) using *BamHI* (NEB) and *EcoRI* (NEB). These constructs were subsequently linearized with *PvuI* (NEB) and transfected into LIGIV-null cells. 48 hr after transfection, the cells were subcultured under limiting

dilution into 96-well plates with 1 mg/ml G418. Colonies were expanded for 4 to 5 weeks and then complemented clones were identified by immunoblotting as described above.

#### **GCR test**

G-banding cytogenetic analyses were performed in the Cytogenetics Core Laboratory at the University of Minnesota.

## Results

### The inactivation of *LIGIV* in the HCT116 cell line

We used rAAV gene targeting method to inactivate the *LIGIV* gene in the HCT116 colorectal carcinoma cell line (334). This cell line has been used extensively for similar gene-targeting studies. The cell line is immortalized, transformed and defective in the *MLH1* gene making it mismatch repair defective (359-362). However, the cell line is diploid, has an unusually stable karyotype and is wild type for almost all other DNA repair and checkpoint genes (359). The *LIGIV* gene is located on chromosome 13 and has three exons. Since the *LIGIV* coding sequence (CDS) resides only on exon 3, we disrupted the first 302 bp of the CDS by replacing it with the neomycin phosphotransferase (NEO) gene. The relevant rAAV targeting vector consisted of two ~900 bp homology arms (which, on the chromosome, flank the 302 bp region to be deleted), and a NEO selection cassette, which was itself flanked by LoxP sites (Fig. 1A).

HCT116 cells were infected with the rAAV<sup>LIGIV</sup> vector and in the first round of targeting, 177 G418-resistant clones were subsequently screened by PCR using primers that were specific to the targeting vector and a unique sequence in the flanking DNA. Two correctly targeted clones (#130 and #163) were identified for a relative gene targeting efficiency of 1.1%, which is very similar to the frequencies reported for other rAAV-mediated gene targeting studies. To ensure that isolated clones originated from a single cell, we subsequently subcloned them. One of the subclones (#130-26) was designated as *LIG4*<sup>+NEO</sup> and then treated transiently with a Cre-recombinase expression vector to remove the NEO selection cassette (which is flanked by LoxP sites; Fig. 1A). A G418-sensitive derivative subclone was obtained from this protocol and renamed as *LIG4*<sup>+/-</sup> or Cre1. Cre1 cells were then subjected to a second round of gene targeting — using the exact same rAAV<sup>LIGIV</sup> vector that had been used in the first round — to inactivate the remaining wild-type allele. In the second round of gene targeting, we screened a total of 673 clones and obtained 16 correctly targeted clones (relative gene targeting frequency: 2.4%). Surprisingly, 15 of these clones were retargeted to the already inactivated allele and were therefore still heterozygous. Only one clone (#312, see below) was targeted at the remaining wild-type allele and was designated as

LIG4<sup>NEO/-</sup>. This clone was subsequently converted into a LIG4<sup>-/-</sup> cell line by transient treatment with the Cre recombinase as described above.

For the second round of screening, we used a primer set that could distinguish the three possible different LIGIV allelic states: a wild-type allele, a NEO-targeted allele and the Cre-treated allele with a residual single LoxP site. These three configurations generated 681 bp, 2595 bp and 449 bp PCR products, respectively (Fig. 1B). Retargeted clones, where the targeting construct integrated into the previously inactivated allele, had a diagnostic pattern where the 449 bp band from the Cre-treated allele disappeared but where the 681 bp band from the wild-type allele was still present (*e.g.*, clones #69 and #198, Fig. 1C). In contrast, randomly-targeted second round clones still retained both the 449 bp and 681 bp bands in a manner unchanged from the parental Cre1 clone (*e.g.*, clone #417 and Cre1, Fig. 1C). Apart from all these clones was the pattern generated by clone #312, which did not produce the 681 bp band from the wild-type allele but did generate the 449 bp band originating from the Cre-treated allele, suggesting that it was a true null clone (clone #312, Fig. 1C). In this PCR analysis, only a 2 min extension time was used, so the 2595 bp product from the targeted allele was not detected (Fig. 1C). Consequently, we further confirmed that the founding LIGIV<sup>+/-</sup> cell lines, Cre1 and #160, have reduced amounts of LIGIV protein and that the LIGIV<sup>-/-</sup> cell line, #312, expressed no LIGIV protein detectable by western blot analysis (Fig. 2). From these molecular and biochemical analyses, we concluded that clone #312 corresponded to a viable LIGIV-null cell line and is hereafter referred to as the LIG4<sup>-/-</sup> cell line.

### **A LIGIV deficiency causes a mild growth defect**

Three thousand cells corresponding to each of the parental (*i.e.*, LIGIV<sup>+/+</sup>), LIGIV<sup>+/-</sup> (Cre1 or #160) or LIG4<sup>-/-</sup> cell lines were seeded on a 6-well plate on day 0 and the growth of cells from day 6 to day 10 was determined. LIGIV<sup>+/-</sup> cells grew at a rate that was slightly, but significantly, reduced from the parental cell line (except on day 10) demonstrating that there is a mild haploinsufficiency associated with LIGIV (Fig. 3). Moreover, LIGIV<sup>-/-</sup> cells grew slower than either of the two LIGIV<sup>+/-</sup> cells lines, especially at earlier time points (Fig. 3). The growth defect associated with the LIGIV-

null cell line, however, in contrast to the growth defect reported for DNA-PK<sub>cs</sub> (363), another key C-NHEJ gene, was quite mild.

### **A *LIGIV* deficiency induces genomic instability**

One of the deleterious consequences of unrepaired DNA DSBs is genomic instability, which predisposes cells to cancer (364). Because *LIGIV* is essential for C-NHEJ activity, we expected that *LIGIV*<sup>-/-</sup> cells might be genomically unstable, and tested this by measuring karyotypic abnormalities using standard G-banding of metaphase chromosomes (Table 1). The background rate for finding detectable GCRs (gross chromosomal rearrangements/abnormalities) is 10% or lower for the parental cell line (115, 363). Twenty metaphases from three independent *LIGIV* heterozygous clones were analyzed: #163-7 and #130-26 are two *LIGIV*<sup>+/-</sup> subclones obtained from the first round of targeting while #69 is one of the second-round retargeted *LIGIV*<sup>+NEO</sup> clones (Fig. 1C). Collectively, the *LIGIV*<sup>+/-</sup> cells showed an average of 10% GCRs (range of 0% to 20%), which was almost the same as that of the parental cell (Table 1), suggesting that there is no haploinsufficiency for genomic instability, in spite of the slower cell growth. In contrast, the *LIGIV*<sup>-/-</sup> cell line had a 30% GCR rate suggesting that *LIGIV* is a suppressor of genomic instability (Table 1).

### ***LIGIV*<sup>-/-</sup> cells are extremely sensitive to DNA damaging agents**

The very first *LIGIV* patient described in the literature developed leukemia, but she actually died from an extreme adverse response to the radiation treatment for this cancer and not from the leukemia itself (238). To experimentally determine if this effect could be recapitulated in our somatic cell model, we carried out colony forming assays in the presence or absence of etoposide, a topoisomerase II inhibitor and a strong radiomimetic drug (365) (Fig. 4). Positive control cell lines included the parental, wild-type cell line and clone #1, which was a randomly targeted clone from the first round of gene targeting and it therefore retained the *LIGIV*<sup>+/+</sup> genotype. As negative controls we used a clone, #70-32, which has a *Ku86*<sup>+/-</sup> genotype and is known to be etoposide-sensitive (366). Compared to the wild-type and randomly targeted clone, the *LIGIV*<sup>+/-</sup> clone showed a mild sensitivity comparable to that of the *Ku86*<sup>+/-</sup> cell line (Fig. 4). In stark contrast, the

LIGIV<sup>-/-</sup> cell line was extremely sensitive to even low concentrations of etoposide (Fig. 4). This result strongly suggests that LIGIV is required for the repair of etoposide-induced DNA DSBs.

### **LIG4<sup>-/-</sup> cells show an increased use of microhomology for DNA end joining**

The A-NHEJ DNA DSB repair pathway is generally only detectable when C-NHEJ is deficient. Indeed, microhomology-mediated end joining (the hallmark repair signature of A-NHEJ) is increased in the 180BR cell line, which was derived from a LIGIV patient (244). To see if these attributes could be extended to our cell line, we tested the DNA end joining activity of LIGIV<sup>-/-</sup> cells using a reporter substrate, pDVG94, that can differentiate C-NHEJ and A-NHEJ products (254). Digestion of pDVG94 with *EcoRV* and *AfeI* restriction enzymes generates a blunt-ended linear substrate that has 6 bp repeats of homologous sequence (ATCAGC) at both ends (Fig. 5A). When this substrate is transfected into mammalian cells it can be joined by either C-NHEJ or A-NHEJ, but only when it is repaired by A-NHEJ will a novel *BstXI* restriction recognition site be generated (Fig. 5A). Consequently, this linear substrate was transfected into cells and recovered 48 hr later. Repaired junctions were amplified by PCR with radiolabeled primers, and the resulting 180 bp radiolabeled PCR products were digested with *BstXI*. The amount of 180 bp uncut product represents the repair carried out by C-NHEJ whereas the 120 bp *BstXI*-digested product corresponds to repair by A-NHEJ. Wild-type cells had less than 1% A-NHEJ activity (Fig. 5B, lane 2), consistent with previous analyses (254). In contrast, more than 99% of the repair in LIGIV<sup>-/-</sup> cells was A-NHEJ-mediated (lane 4). Based upon these results we concluded that LIGIV is the major C-NHEJ pathway ligase and in the absence of LIGIV cells use the A-NHEJ pathway for DSB repair.

### **Complementation experiments using the re-expression of a wild-type LIGIV cDNA**

To confirm that the phenotype we observed in LIG4<sup>-/-</sup> cells was due to the absence of LIGIV, we attempted to rescue the LIG4<sup>-/-</sup> cells using the expression of a LIGIV cDNA (Fig. 6). A wild-type LIGIV cDNA was cloned into the pcDNA3.1(+) mammalian expression vector with or without an HA epitope tag. These constructs were transfected into the LIG4<sup>-/-</sup> cells and G418-resistant colonies were screened by immunoblot analysis

for the expression of *LIGIV* as described in the Materials and Methods section. Four clones, two each for each of the cDNAs, were used for further characterization. Clones #14 and #57 expressed untagged wild-type *LIGIV* while clones #17 and #52 were complemented with the HA epitope-tagged wild-type *LIGIV*. Clones #14 and #52 showed higher levels of expression than clones #57 and #17, respectively (Fig. 6). First, we tested the DNA end joining activity of the complemented clones. All four clones showed a significant shift towards the wild-type profile (Fig. 5B, lanes 6, 8, 10 and 12). The clones with the lower level of *LIGIV* expression (clone #57 and #17; Fig. 6) also showed the lower degree of repair complementation (Figure 5B, lane 8 and 10). We also tested the etoposide sensitivity of the complemented clones. Again, two of the clones (clones #14 and #52) showed better (albeit not wild-type levels) of complementation than clone #57 and #17, respectively (Fig. 7). The ability to complement the repair and etoposide-sensitivity defects of our *LIGIV*-null cell line by the re-expression of *LIGIV* strongly suggests that the phenotypes of this cell line are due specifically to the loss-of-function of *LIGIV*.

## Discussion

In this study, we generated a viable  $LIGIV^{-/-}$  HCT116 cell line using rAAV-mediated gene targeting methodology. The inactivation of  $LIGIV$  in two independent human cell lines, NALM-6 (207) and HCT116 {this study}, demonstrates that  $LIGIV$  is not essential for human somatic cell viability. With this said, it is important to note that multiple independent studies using rAAV-mediated gene knockout strategies have demonstrated that there is no allelic preference in gene targeting (367-370). However, in the second round of  $LIGIV$  gene targeting, we observed a large bias toward retargeting at the already-inactivated locus compared to targeting of the wild-type locus, with 15 of the former events and only one of the latter. This disequilibrium in the gene targeting frequency is usually indicative of genes that provide a significant growth disadvantage when absent. This is the case for  $DNA-PK_{cs}$  (363) and  $PARP-1$  {M. Mueller & E. A. Hendrickson, unpublished data} where highly skewed second round gene targeting frequencies (1 out of 17 and 2 out of 32 correctly targeted clones for  $DNA-PK_{cs}$  and  $PARP-1$ , respectively) were observed. Not surprisingly, the resulting  $DNA-PK_{cs}$ -null and  $PARP-1$ -null cell lines showed very severe proliferation defects with cell doubling times on the order of ~40 hr {(363); M. Mueller & E. A. Hendrickson, unpublished data}. In contrast, the  $LIGIV$ -null cell line reported here, while it does display growth defects (Fig. 3), still doubles at a fairly robust rate. The discrepancies in these studies may be related to genomic stability. The loss of both alleles of  $DNA-PK_{cs}$  for example, results in a cell line with a very high rate of GCRs in which 75% of all metaphases show at least one gross karyotypic abnormality (363). This level of genomic instability is certainly deleterious and likely explains the large deficits in proliferation of the  $DNA-PK_{cs}$ -null cells.  $LIGIV$ -null cells, on the other hand, had a much lower (albeit elevated) frequency of GCRs (30%; Table 1). From these considerations we conclude that while  $LIGIV$  is unequivocally non-essential for human somatic cells, it nonetheless provides important cellular functions such that cells without  $LIGIV$  are at a distinct disadvantage for growth and survival in comparison to  $LIGIV$ -proficient cells.

### **C-NHEJ is not essential for human cell viability**



We have demonstrated that LIGIV is not essential for human somatic cell viability. Since LIGIV is absolutely required for C-NHEJ, it would be logical to extrapolate our results to conclude that C-NHEJ is also non-essential in human somatic cells. Confusing this interpretation, however, is the observation that two other major C-NHEJ components, Ku70 and Ku86, are essential in human somatic cells (115, 117, 276). We believe this discrepancy is due to the fact that LIGIV likely functions exclusively in C-NHEJ, whereas Ku has additional (essential) roles in telomere maintenance (139, 371). Unlike LIGIV, Ku70 and Ku86 are physically associated with telomeres (112, 113, 119), and once either Ku subunit is depleted, the telomeres shorten and then fuse (119). Previous research from our laboratory showed that a Ku86 deficiency causes massive telomere loss, sister chromatid fusion and t-circle formation, which eventually induces cell death within a week (117). To investigate whether LIGIV is involved in telomere fusion under Ku86 deficient conditions, we have generated a Ku86 and LIGIV doubly-null cell line {Y. Wang and E. A. Hendrickson, unpublished data}. This doubly-null cell line still displays telomere fusions {Y. Wang and E. A. Hendrickson, unpublished data}, implying that only Ku86, but not LIGIV nor C-NHEJ, is involved in telomere maintenance in human cell. Consistent with these observations are the recent results demonstrating that the functional inactivation of the two LIGIV accessory factors, XLF {F. Fattah and E. A. Hendrickson, manuscript in preparation} and XRCC4 {B. Ruis and E. A. Hendrickson, manuscript in preparation} result in viable cell lines. Thus, the preponderance of data suggests that not only is LIGIV not essential in human somatic cells, but that C-NHEJ is non-essential as well.

### **LIGIV is nonetheless probably essential for early development**

To date, only seven LIGIV-defective patients have been reported in a world-wide population of seven billion (230, 240, 243). This paucity of patients is analogous to the gene targeting disequilibrium we observed and suggests strongly that there is a profound selective disadvantage to having reduced levels of, or being totally without, LIGIV. Four of the LIGIV-defective patients presented with a “LIGIV syndrome”, which is associated with chromosomal instability, pancytopenia, developmental and growth delay and dysmorphic facial features (239). Two other patients had leukemia (*i.e.*, cancer

predisposition) (238, 241) and the last patient presented as a T<sup>B</sup>NK<sup>+</sup> RS-SCID (243). The LIGIV mutations in all of these patients have been characterized at least at the DNA level. One patient had a mutation in the putative nuclear localization signal (NLS), so that the protein, which was otherwise presumably functional, was mislocalized in the cytoplasm. Other patients had unique kinds of homozygous or compound heterozygous hypomorphic mutations, which did not completely abolish, but significantly reduced, enzyme function (242), and the severity of the clinical features was correlated with the level of residual LIGIV activity (241). In summary, even though LIGIV is dispensable for human somatic cell survival, the absence of unequivocally LIGIV-null patient suggests that there is an essential role for LIGIV somewhere during early development. This stage may correspond to neural development because two of the main clinical features of LIGIV patients are microcephaly and neurological abnormalities (239). This view is consistent with work carried out in mice. LIGIV is essential in the mouse and the mice succumbed very early in development not to any obvious defect in C-NHEJ, but due to massive neuronal cell death (209, 210). The extraordinary requirement for LIGIV during neural development is not obvious. The brain does, however, occupy only 2% of the total body weight, but consumes 20% of the cellular oxygen (372-374). Perhaps this high oxidative stress causes more DNA DSBs in brain tissue compared to other tissues/organs, which would clearly be detrimental to cells without functional LIGIV. This speculation is at least consistent with the extreme sensitivity of LIGIV-null cell to DNA damaging agents (Fig. 4), where cells encountering high levels of DNA DSBs may resemble cells required for early neuronal development subjected to high oxidative stress. In summary, we believe that the data are compelling that LIGIV is likely essential for the development of an organism, but dispensable for the survival of single cells.

### **Ku, and not LIGIV, is the main regulator of repair pathway choice**

The viability of LIGIV-null cells suggests that other DNA DSB pathways (*i.e.*, HR and A-NHEJ) are sufficient to deal with the endogenous DNA DSB damage that arises in normal cells. This observation, however, begs the question of how a normal cell that suffers a DNA DSB decides which DSB repair pathway to use to enact repair. This issue is quite important since the repaired products these pathways generate are distinctly

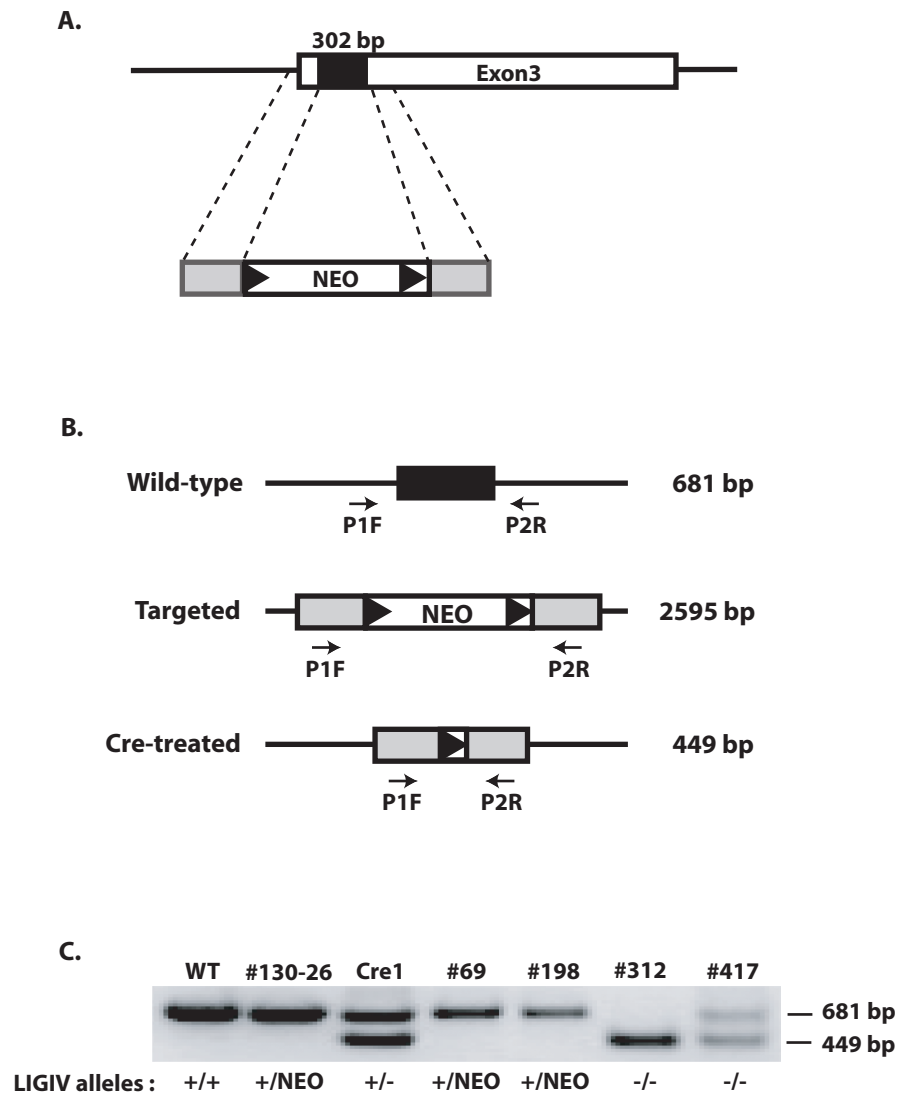
different. Numerous recent studies have suggested that this regulatory activity resides with Ku (19, 27, 254). C-NHEJ, the major DNA DSB pathway, is usually depicted as a sequential event with the first protein in the reaction mechanism, Ku, essentially acting to commit the DSB to C-NHEJ while simultaneously suppressing the access of HR and A-NHEJ factors to the DSB (254). Our work is consistent with that view. Thus, in a recent study from our laboratory our *LIGIV*-null cell line was shown to have lost virtually all of its end joining activity (254). That study and our work here demonstrate that the residual end joining activity that remains in *LIGIV*-null cells corresponds almost exclusively to A-NHEJ (Fig. 5). The phenotype of *LIGIV*-null cells stands in stark contrast to that of Ku-reduced cells, which — while having no detectable C-NHEJ activity — actually have elevated levels of HR (158, 227, 249) and A-NHEJ (244, 254, 347). Nonetheless, it may be possible that *LIGIV* retains a subtle role for pathway choice regulation although that is not obvious from our studies. Thus, *LIGIV:XRCC4* has been reported to be required for C-NHEJ initiation by facilitating polymerase and nuclease activity in HeLa cell extract (226, 227). Moreover, in yeast, *Dnl4* and *Lif1*, yeast orthologs of *LIGIV* and *XRCC1*, are required to stabilize NHEJ complexes and suppress HR by inhibiting resection (227). Combined, these results still leave the possibility that initiation and/or stabilization of C-NHEJ by *LIGIV* may be required to suppress A-NHEJ, although that was not so evident in our studies.

Interestingly, it could also be argued that *LIGIV* has a small role in suppressing HR. Thus, the first round of gene targeting for *LIGIV* was 1.1% whereas the second round gene targeting frequency was 2.4%, suggesting that reduced *LIGIV* levels enhance HR. Moreover, in additional gene targeting studies, we have shown that the targeting frequency at the *Ku70* locus in *LIGIV*-null cells is 2-fold elevated compared to wild-type cells (Supplementary Table 1; F. Fattah and E. Hendrickson, unpublished data). Similarly, the targeting frequency at the *RAD54B* locus is increased 2-fold in *LIGIV*-null cells (Supplementary Table 1; unpublished data). Together, these studies show conclusively that in the absence of *LIGIV*, gene targeting — clearly an HR-mediated reaction — is slightly enhanced. Nonetheless, this enhancement is nowhere near the improved gene targeting frequencies reported for Ku-reduced human cell lines (276). In summary, our data are consistent with a mechanism where the bulk of the pathway choice

regulation resides within the purview of Ku, but that the presence or absence of LIGIV may influence these processes as well.

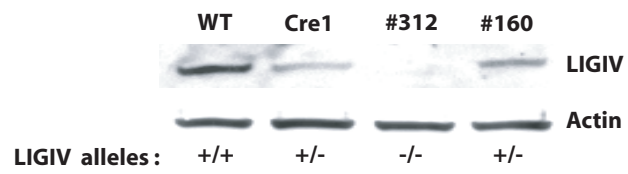
**Figure 1.** The targeting strategy and primers used for the LIGIV knockout. (A) The black rectangle represents the first 302 bp of LIGIV coding sequence, which was replaced by the NEO gene using rAAV-mediated gene targeting. The gray rectangles represent the ~900 bp of homology arms flanking the targeting region. The black triangles represent LoxP sites. (B) Primers used for the second round of screening. The P1F/P2R primer set gives different-sized PCR products depending on the allelic states of the LIGIV gene; 681 bp for the wild-type allele, 2595 bp for the NEO-targeted allele, and 449 bp for a cre-treated allele. Under the PCR condition we used here with short extension time, the 2595 bp product was not generated. (C) PCR screening. #130-26 is one of the LIGIV<sup>+NEO</sup> subclones. Cre1 is a G418-sensitive LIGIV heterozygous cell generated by Cre-recombination of #130-26. Cre1 was used for the second round of targeting. Cell lines #69 and #198 are retargeted clones from the second round of gene targeting. Clone #312 is LIGIV<sup>-/-</sup> and #417 is an example of a LIGIV<sup>+/-</sup> clone that suffered a random integration during the 2<sup>nd</sup> round of gene targeting.

**Figure 1. Targeting construct and primers used for LIGIV gene targeting**



**Figure 2.** Inactivation of the LIGIV gene is confirmed at the protein level. Western blot analysis confirms that clone #312 is LIGIV<sup>-/-</sup> cell. Cre1 is a G418-sensitive LIGIV heterozygous cell and clone #160 is a LIGIV<sup>+/-</sup> clone that suffered a random integration during the 2<sup>nd</sup> round of gene targeting.

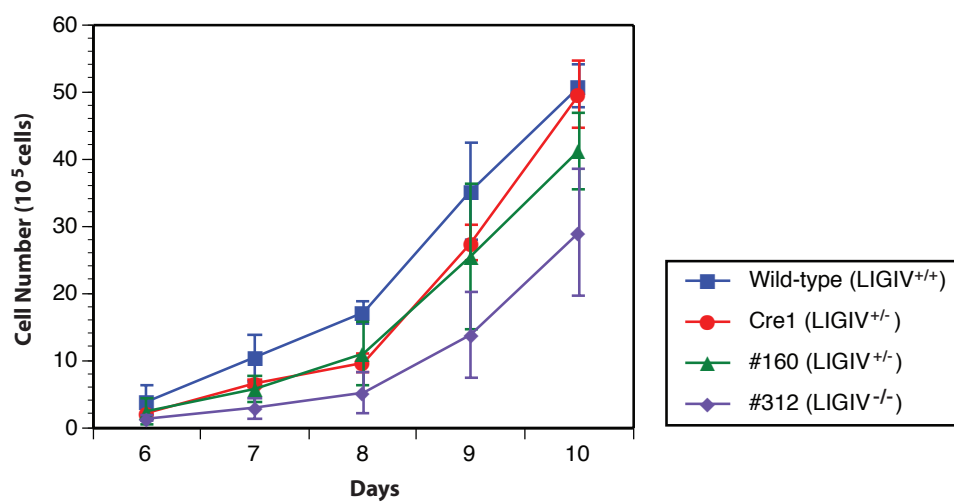
**Figure 2. Immunoblotting analysis of *LIGIV* targeted clones**





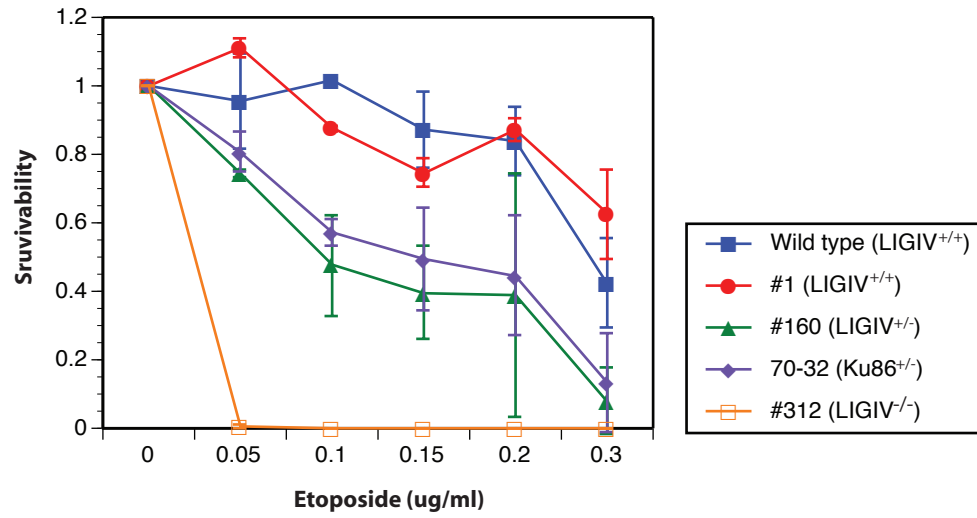
**Figure 3.** The  $LIGIV^{-/-}$  cell line has a mild growth defect. On day 0, three thousand cells were seeded into 6-well plates and the cells were counted from day 6 to day 10. Cre1 and #160 are  $LIGIV^{+/-}$  clones. The data shown here represent the average of six counts that came from four independent sets of duplicates excluding the maximum and minimum values.

**Figure 3. LIGIV deficiency causes mild growth retardation**



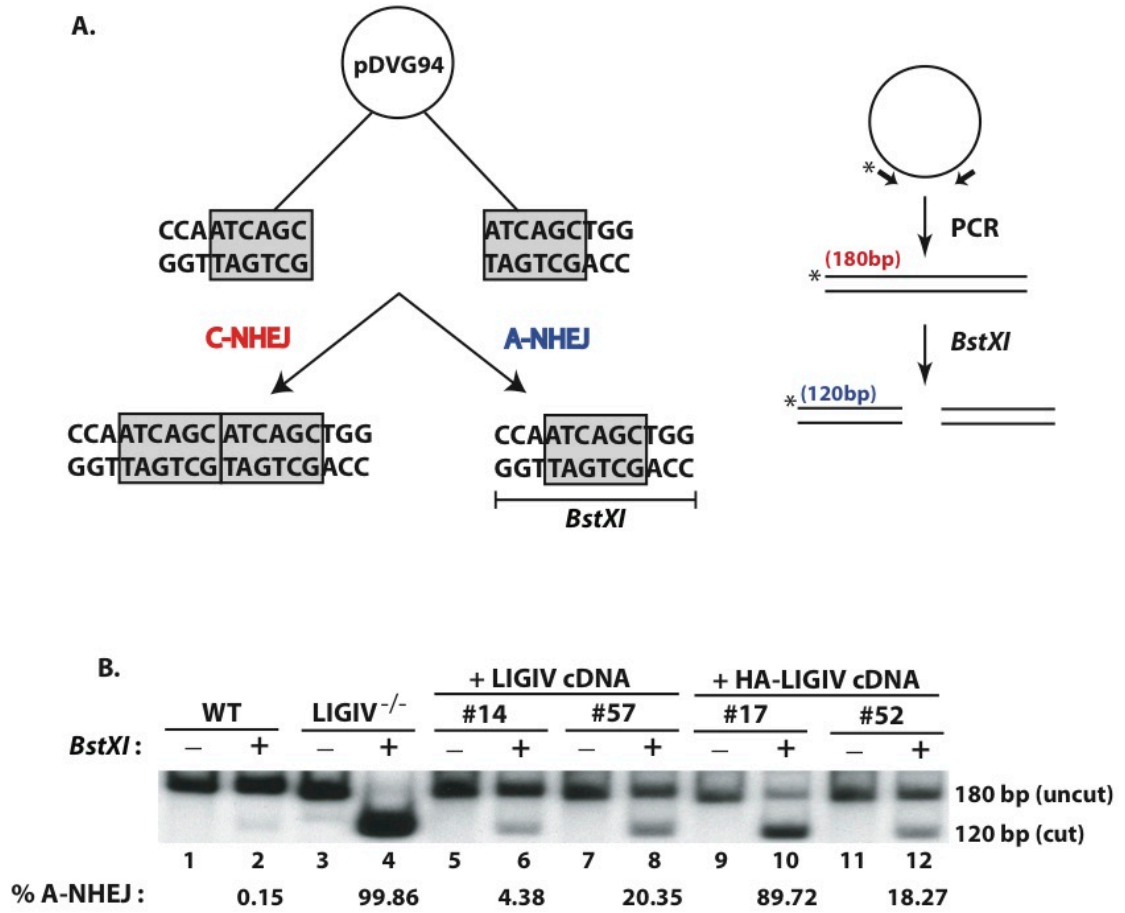
**Figure 4.**  $LIGIV^{-/-}$  cells are extremely sensitive to DNA damaging agents. On day 0, three hundred cells were subcultured into 6-well plates with different concentrations of etoposide. 10 to 14 days later, the cells were fixed and stained with crystal violet. The number of colonies that survive at different concentrations of etoposide was normalized by the survivability of untreated cells, which was set as 1. Clone #1 is one of the first-round random integrates, so it still has wild-type  $LIGIV$  alleles. Clone 70-32 is an etoposide-sensitive  $Ku86^{+/-}$  HCT116 cell line. Each value represents the average of two independent experiments.

Figure 4. LIGIV null cell displays extreme sensitivity against etoposide



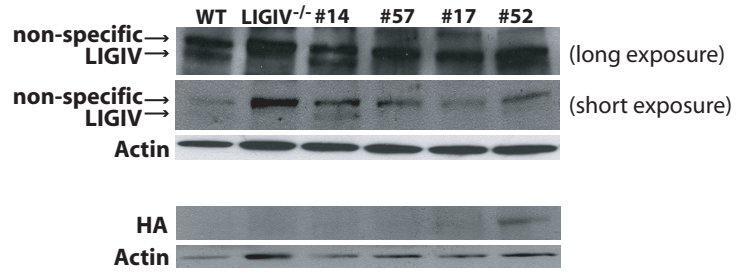
**Figure 5.** *LIGIV*<sup>-/-</sup> cells use microhomology for DSB repair. (A) After *EcoRV* and *AfeI* restriction enzyme digestion, the reporter substrate, pDVG94, becomes a blunt-ended linear plasmid with 6 bp direct repeats at both ends. Repair of the plasmid via C-NHEJ retains part of both repeats, whereas A-NHEJ generates only a single repeat, which can be subsequently cleaved by *BstXI*. Repaired junctions can then be amplified by PCR using radiolabeled primers and the 180 bp PCR product is subjected to *BstXI* digestion. The 180 bp uncut product represents repair via C-NHEJ whereas the 120 bp digested product represents A-NHEJ-mediated repair. (B) In *LIGIV*<sup>-/-</sup> cells, virtually almost all of the repair is mediated by A-NHEJ (lane 4). After complementation (lanes 6,8,10, and 12), the ratio between the 180 bp versus 120 bp products becomes more similar to that of wild-type (lane 2), indicating the restoration of functional C-NHEJ activity in the complemented cells.

Figure 5. Microhomology mediated DNA end-joining activity in LIGIV-deficient cells



**Figure 6.**  $LIGIV^{-/-}$  complementation. Clones #14 and #57 are complemented with an untagged wild-type  $LIGIV$  cDNA. Clone #17 and #52 are complemented with a HA-tagged wild-type  $LIGIV$ . In the  $LIGIV$  immunoblots, there is a non-specific band right above the authentic  $LIGIV$  band. Both a short exposure and a long exposure of the same blot are presented. Clones #14 and #52 express more  $LIGIV$  protein than clones #57 and #17, respectively. The same samples were blotted with a HA-antibody as well. Actin has used as a loading control.

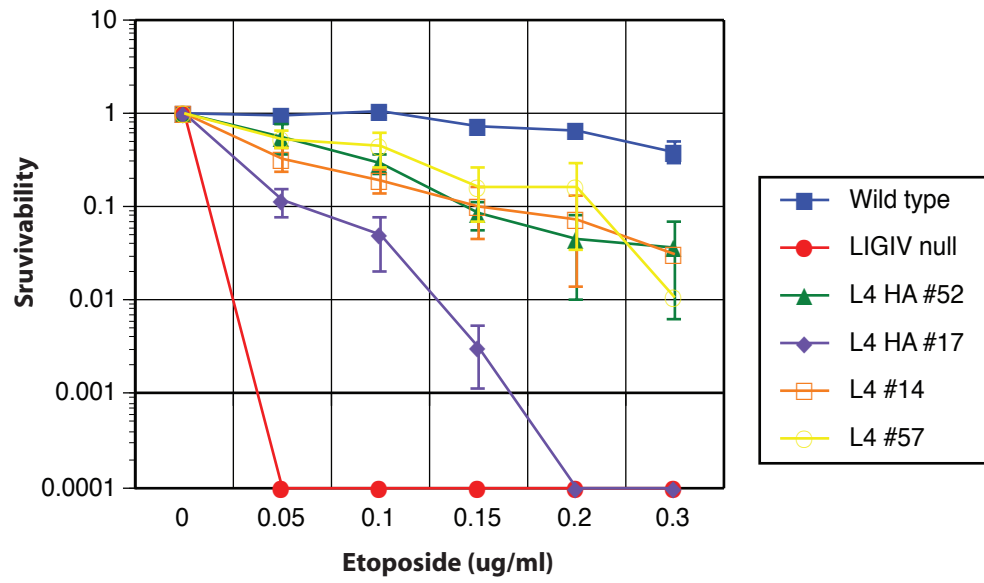
**Figure 6. Immunoblot result of LIGIV complemented clones**





**Figure 7.** Complemented clones are more resistant to etoposide than  $LIGIV^{-/-}$  cells. The etoposide sensitivity test was performed exactly as described in Figure 4.

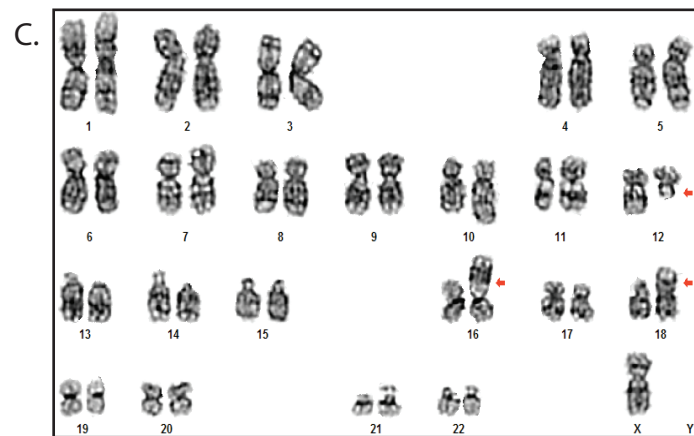
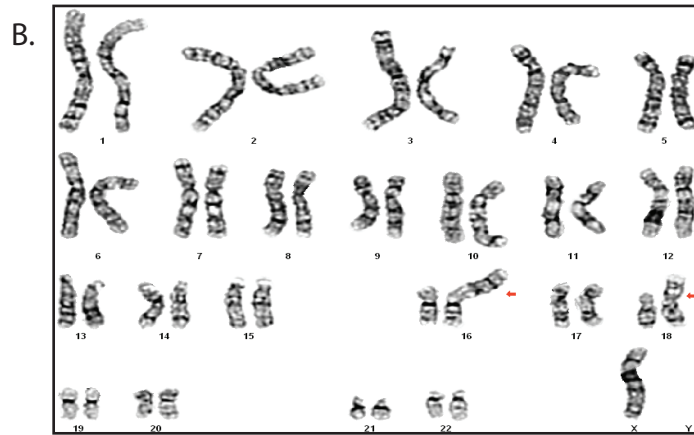
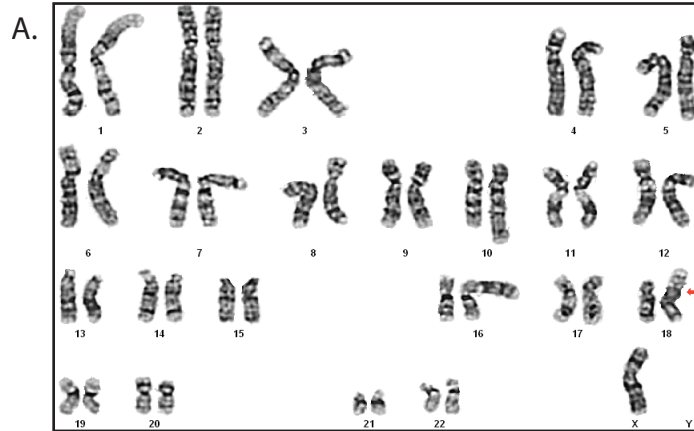
Figure 7. LIGIV complementation rescues extreme etoposide-sensitivity of LIGIV null cell



**Table 1. Summary of GCR frequency in *LIGIV* clones**

<b>Designation</b>	<b>Genotype</b>	<b>Metaphases</b>	<b>GCRs</b>	<b>Frequency (%)</b>
<b>wtHCT116</b>	<b><i>LIGIV</i><sup>+/+</sup></b>	<b>143</b>	<b>14</b>	<b>9.8</b>
<b>#69</b>	<b><i>LIGIV</i><sup>+/-</sup></b>	<b>20</b>	<b>0</b>	<b>0.0</b>
<b>#130-26</b>	<b><i>LIGIV</i><sup>+/-</sup></b>	<b>20</b>	<b>2</b>	<b>10.0</b>
<b>#163-7</b>	<b><i>LIGIV</i><sup>+/-</sup></b>	<b>20</b>	<b>4</b>	<b>20.0</b>
<b>#312</b>	<b><i>LIGIV</i><sup>-/-</sup></b>	<b>20</b>	<b>6</b>	<b>30.0</b>

**Supplementary Figure 1.** G-band karyotype. The red arrows show instances of observed aberrant chromosomal abnormalities. (A)  $LIGIV^{+/-}$  #163-7. (B)  $LIGIV^{+/-}$  #130-26. (C)  $LIGIV^{-/-}$  #312.



**Supplementary Table 1. Gene targeting rate in LIG4 deficient cell lines**

Targeted Genes Cell Lines	LIG4	Rad54B	Ku70
wtHCT116	1.13	*n.a.	1st : 0.69 2nd : 4.12
LIG4 <sup>+/-</sup>	2.16	*n.a.	0.73
LIG4 <sup>-/-</sup>	*n.a.	1st : 8.82 2nd : 13.19 3rd : 4.74	2.20

\*n.a. = not available

## **CHAPTER III**

**LIGIII is an essential gene for the survival of human  
somatic cells**

C-NHEJ (classic non-homologous end joining) is the primary DNA DSB (double-strand break) repair pathway in higher eukaryotes. Recently, evidence has been accumulated for the existence of a novel, albeit minor, end joining repair pathway. This pathway can normally be detected only under C-NHEJ deficient condition and is called A-NHEJ (alternative-NHEJ). Currently, the molecular mechanism of A-NHEJ is very obscure. Nonetheless, several proteins have been implicated in the A-NHEJ pathway including LIGIII:XRCC1, PARP1, MRN, CtIP, and Histone H1. Among these genes, a dependency on LIGIII is one of the most prominent features of A-NHEJ. To experimentally test the requirement for LIGIII in A-NHEJ we constructed a LIGIII conditionally null human cell line using rAAV (recombinant adeno-associated virus) – mediated gene targeting methodology. Our data demonstrate that, similar to studies recently described in the mouse and chicken cell line, human LIGIII is required for mitochondrial function and this defines its essential activity for cell viability. Disappointingly, however, DNA repair activity in the nucleus appeared utterly unaffected by the deficiency in LIGIII. Moreover, A-NHEJ was predicted to be the primary mediator of rAAV random integration events, but the same rAAV-mediated gene targeting rate was observed in LIGIII-null cells as in wild-type cells. In conclusion, we have shown that human LIGIII has an essential function in mitochondria maintenance, but we find that it is dispensable for most types of nuclear DSB repair.



## Introduction

A DNA DSB is one of the main challenges that undermine genomic integrity (2). In order to avoid the destructive and pathological outcomes that result from unrepaired DNA DSBs, it is crucial to have efficient DSB repair pathways. In most organisms, there are two major DSB repair pathways, HR (homologous recombination) and C-NHEJ (10, 18). HR is preferentially used in lower organisms, however, in mammalian cells — and particularly in human cells — the majority of DSBs are repaired via the C-NHEJ pathway where Ku70:Ku86 and LIGIV:XRCC4 are the major players. In C-NHEJ, DSB ends are bound by Ku70:Ku86 heterodimer, which subsequently recruits DNA-PK<sub>cs</sub>, the catalytic subunit of the DNA-PK holoenzyme (131, 132). A structure-specific nuclease, Artemis, cleans up DSB ends that aren't flush or overlapping, and these polished broken ends are then covalently rejoined by a ligation complex, composed of LIGIV, XRCC4, and XLF (190, 204, 237).

Recently, it has become apparent that there is a minor end-joining pathway present in higher eukaryotes that is generally only detectable in the absence of C-NHEJ. It has interchangeably been referred to as MMEJ (micro-homology-mediated end joining), B-NHEJ (backup-NHEJ) and A-NHEJ, but for the purposes of consistency, we will restrict ourselves to the use of just the latter name. The first description of A-NHEJ came from an analysis of the chicken DT40 cell line (159). In chicken cells (as well as in all mammalian cells) DSB repair is always biphasic, with fast (< 30') and slow (4 to 24 hr) components. When mutations were made in DT40 C-NHEJ genes, the fast component of DSB repair completely disappeared, but the slow component did not. The authors went on to show that the slow component was not due to HR and thus they surmised that it must be some sort of novel end joining, *aka* A-NHEJ (159). Confirmation of the existence of A-NHEJ came during an analysis of the 180BR cell line, a primary human fibroblast cell line containing inactivating mutations in LIGIV (244). Interestingly, even without functional LIGIV, the ligase known to be essential for C-NHEJ, 180BR cell were still able to repair some DSBs (244); again implying the existence of an alternative end joining pathway.

The mechanism of A-NHEJ was initially (and actually, for the large part still is) obscure. It is presumed that some other end-binding factor besides Ku is required to bind

onto the ends and funnel the DSBs into the A-NHEJ pathway. Then, because microhomology appears to be used at high frequency to mediate the repair event, some end resection (perhaps extensively so) is required. Alignment activities to bring the microhomologies into register are probably 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. Because the pathway uses microhomology to mediate the repair event, deletions (perhaps extensive ones) always accompany the repair event as does loss of one of the blocks of microhomology (274, 375).

With interest in A-NHEJ sparked, several laboratories made dedicated genetic and biochemical attempts to identify the factors associated with the pathway. In particular, a brute force nuclear extract fractionation protocol identified LIGIII, heretofore known only for its role in BER (base excision repair), as the ligase required for A-NHEJ (194). Using guilt by association as a scientific rationale, PARP1 and XRCC1, two proteins known to interact with LIGIII during BER, were subsequently identified as also being involved in A-NHEJ (27, 376). PARP1 is presumed to compete with Ku for binding to broken DNA ends thereby dictating pathway choice whereas XRCC1 appears to act as a chaperone or stability function for LIGIII (247, 248, 377). Additional factors have also been implicated in A-NHEJ. A recent report suggests that histone H1 may act as an alignment factor — an important activity given the role of microhomology needed to mediate the repair event — during A-NHEJ (250). CtIP and the MRN complex — factors known to be involved in the end resection events required for HR — have, not surprisingly, (given that most of the junctions repaired by A-NHEJ possess 5 to 25 nt of microhomology) also been implicated in the end resection steps of A-NHEJ (22, 43, 44, 252).

If the gene factors that are needed for A-NHEJ are still not completely defined and the A-NHEJ reaction mechanism is still somewhat nebulous, it is probably fair to say that the biological role(s) of A-NHEJ is even more poorly understood. Interest in A-NHEJ was originally piqued, when it was demonstrated to provide a back-up mechanism for CSR (class switch recombination). CSR is a lymphoid-restricted, quasi site-specific recombination that is required for the isotype switching of antibodies; a process needed for the primary antibody that can successfully engage other effector functions (4). CSR,

like V(D)J recombination, was long considered to be an exclusively C-NHEJ-dependent process. Later work, however, demonstrated that significant amounts of CSR could still be detected even when C-NHEJ had been ablated and this activity was then shown to be due to A-NHEJ (347-349). 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 shown that most of these chromosomal translocations have extensive microhomology at their breakpoint junctions, which strongly implicates A-NHEJ in their genesis (378-380). This hypothesis has gained support from a recent report from the Jasin group, where LIGIII conditionally-null mice cells showed decreased translocation frequencies and reduced microhomology usage (269, 272). Finally, our laboratory has demonstrated that a reduction in Ku70 levels increased the rAAV-mediated gene-targeting rate (276). Surprisingly, however, the random rAAV integration rate did not change, which implied that these events may be mediated instead by A-NHEJ (276, 378).

Unlike the HR and C-NHEJ pathways, which are evolutionarily conserved from bacteria to man, the LIGIII-dependent A-NHEJ pathway appears to have evolved quite recently and has only been detected in *Xenopus laevis* and mammals (190). This makes the experimental analysis of A-NHEJ somewhat difficult since the commonly used model systems (*e.g.*, yeast) are not directly relevant. Additionally, some of the accepted A-NHEJ factors are multifunctional and/or essential. For example, PARP1 is known to interact with and PARylate proteins involved in DNA repair, transcription, DNA methylation, and the regulation of chromatin structure and histone modification to control physiological and pathological outcomes (381). Similarly, MRN and CtIP are major components in the end processing required for HR and MRN complex at least also functions as a damage sensor and a signal transducer during DSB repair in general (43, 49, 54, 84). In addition, before it was recognized as a crucial A-NHEJ component, the LIGIII:XRCC1 complex had been studied extensively for its roles in single-strand break repair, short patch base excision repair and nucleotide excision repair (190, 247, 382). This multifunctionality certainly makes biochemical and genetic analyses difficult. Compounding this issue is that fact that functional inactivation of either LIGIII or

XRCC1 in the mouse results in early embryonic lethality (265, 266, 269). Similarly, histone H1 is an essential component of general chromatin (383).

Despite these daunting experimental hurdles, some features of the A-NHEJ pathway have slowly come into focus. First, A-NHEJ-mediated repair events appear to require 5 to 25 nt of microhomology, which is clearly distinguishable from HR or C-NHEJ, where either extensive (50 to hundreds of nt) homology or no homology are required, respectively. Second, the repaired products always have associated deletions. End resection is almost certainly required to find the microhomology near the break site, and once annealing and repair occur, the flap structures are likely degraded, resulting in deletions. The third, and the most distinct feature, is A-NHEJ's LIGIII-dependency. There are three ligases in mammals; LIGI, LIGIII, and LIGIV. LIG1 is required for Okazaki fragment ligation during DNA replication, and LIGIV functions exclusively in C-NHEJ. In contrast, and although experimental proof is still incomplete, LIGIII appears to be the only ligase utilized by the A-NHEJ pathway.

Unlike other ligases, LIGIII is molecularly heterogeneous (261, 262). Alternative splicing generates an ubiquitously expressed *alpha* form and a germ cell-specific *beta* form. Only the mitotic *alpha* form has a C-terminal BRCT domain, which permits interaction with its binding partner, XRCC1 (245, 247, 263). In addition to alternative splicing, alternative translation initiation generates mitochondrial and nuclear forms of LIGIII, which either contain or lack a MLS (mitochondria leading sequence), respectively (262). This heterogeneity of LIGIII molecular isoforms implies diverse functional roles for LIGIII. One experimental approach to unraveling the complexity of LIGIII is to generate a LIGIII-deficient model system. This has already been accomplished in the mouse. LIGIII-null mice were early embryonic lethal (266, 269). Recently, a LIGIII conditionally-null mouse was described that showed significant mitochondrial function deficits, but little in the way of DNA repair phenotypes (269). In contrast, to date there have been neither human LIGIII patients nor LIGIII-deficient human cell systems described.

In this study, we have conditionally inactivated the LIGIII gene in the HCT116 human colorectal carcinoma cell line to generate an A-NHEJ-defective human model system. Based on the previous mice studies, we anticipated that LIGIII would be an

essential gene due to its mitochondrial function (269, 271). Indeed, we were able to confirm this hypothesis and demonstrate that the loss of LIGIII from human cells results in death due to mitochondrial dysfunction. To investigate LIGIII's nuclear A-NHEJ activity, we also constructed a cell line that does not express LIGIII in the nucleus but is kept alive by complementation with a mitochondrial-only form of LIGIII. A nuclear LIGIII deficiency caused mild growth retardation, but disappointingly did not affect the overall repair activity nor sensitivity against DNA damaging agents of the cells. Perhaps in keeping with this finding, we were also able to demonstrate that LIGIII-dependent A-NHEJ does not mediate rAAV random gene integrations.

## Materials and Methods

### Construction of LIGIII targeting vectors

Conditional and regular knockout LIGIII targeting vectors were constructed utilizing the system described, with a few modifications (334). For the conditional knockout vector, the left and right homology arms, the latter of which contained the loxP-flanked exon 4 fragments, were generated by PCR. The primer sets used to make these fragments were: Exon4\_KpnF1/XhoR1 was used to generate the LoxP-flanked exon 4 fragment. For the left homology arm, Exon4\_LARM\_F1 and LARM\_SacR1 primers were used. For the right homology arm, Exon4\_RARM\_XhoF1 and R1 primers were used. The relevant homology arms and a neomycin-resistance gene cassette were assembled together as a fusion PCR product, which was then cloned into the pAAV-MCS vector. A regular knockout targeting vector was generated in a similar way, but it did not include the loxP-flanked exon 4 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 Exon4\_SC\_F2 and NeoR2 primers, and retargeted clones were confirmed by LIG3\_LArm\_F3 and LIG3\_RArm\_R2 primers.

Exon4\_KpnF1:

5'-CCGGTACCGTAGAGATGGGGTCTTTCTTTGTTGC-3'

Exon4\_XhoR1:

5'-CGCTCGAG***ATAACTTCGTATAATGTATGCTATACGAAGTTATCCA***  
GGAGAGACAGAGGGGGCAAG-3' \****Bold Italic*** indicates the LoxP sequence

Exon4\_LARM\_F1:

5'-ACATAAGCGGCCGCAGAGCACTTTGGCATCTGTCTTC-3'

Exon4\_LARM\_SacR1:

5'-GGCGGCCCGCGGAAAAAATTAAAAAATTAGCTGG-3'

Exon4\_RARM\_XhoF1:

5'-CGCTCGAGGGCTTTTATTCTGGACTCTTTTTTTC-3'

Exon4\_RARM\_R1:

5'-ACATAAGCGGCCGCTGGAGTAGGCAAGAGACTCATAC-3'

Exon4\_SC\_F2:

5'-ATGAGCATCCTGAATAGGCCTTTCCTCCGG-3'

NeoR2:

5'-AAAGCGCCTCCCCTACCCGGTAGGGCG-3'

LIG3\_LArm\_F3:

5'-TGCCACCATGTCCAGCTAA-3'

LIG3\_RArm\_R2:

5'-GAGTCCAGAATAAAAGCC-3'

### **LIGIII complementation**

To construct the mitochondrial-only LIGIII cDNA, the second and third ATGs in the ORF (open reading frame) of the LIGIII cDNA were mutated to ATC, to ensure that the nuclear form of LIGIII could not be expressed. This mitochondrial-only LIGIII cDNA was cloned into the pcDNA3.1(+) vector with a C-terminal HA epitope tag. For the nuclear-only LIGIII, 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 LIGIII cDNA was cloned into a modified pcDNA3.1(+), where the standard neomycin-resistance gene had been replaced with a puromycin-resistance gene. Complementation constructs were linearized by *PvuI* (NEB) and transfected with Lipofectamine 2000 (Invitrogen). For selection, 1 mg/ml G418 and 2 µg/ml puromycin, respectively, were used.

### **Immunoblotting**

Whole cell extracts were prepared with RIPA buffer and 30 µg of protein was electrophoresed on 4% to 20% SDS (sodium dodecyl sulfate)-polyacrylamide gel (BioRad) and rabbit, anti-human DNA LigIII antibody (GeneTex) was used at a 1:1000 dilution. Both the HA antibody (Covance) and FLAG antibody (Sigma) were also used at a 1:1000 dilution. The actin antibody (SantaCruz), which was used for the loading controls, was diluted 1:250.

### **IHC (Immunohistochemistry)**

Cells were plated on multi-chamber slides (Thermo Fisher Scientific) a day before analysis and subsequently fixed with 4% (v/v) paraformaldehyde in PBS 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 done with 0.1% Triton X-100. The LIGIII antibody (GeneTex) was used at a 1:1000 dilution and the Alexa Fluor 488 goat, anti-mouse IgG antibody (Invitrogen) was used to visualize LIGIII. DAPI (0.2 µg/ml) was used to stain the nucleus.

### **Etoposide, MMS (methyl methane sulfonate) and CPT (camptothecin) sensitivity**

We performed an etoposide sensitivity assay as described with slight modifications (358). The cells were plated on a 6-well cell culture dish approximately 17 to 19 hr prior to drug treatment. Etoposide was dissolved in DMSO (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. CPT was dissolved in DMSO to give a 10 mM stock solution. Cells were incubated in CPT-containing media for 24 hr and then switched to drug-free media for 7 to 10 days, or until visible colonies (> 50 cells) had formed.

### **DNA end joining assay**

A DNA end joining assay was performed as described (23). Cells were subcultured into 6-well cell culture dishes a day before transfection. pDVG94 plasmid (2.5 µg) digested with *EcoRV* (NEB) and *AfeI* (NEB) was transfected using Lipofectamine 2000 (Invitrogen). Plasmid DNA was recovered using a modified Qiagen mini-preparation protocol at 48 hr after transfection. Repaired DNA junctions were PCR amplified using FM30 and 5'-radiolabeled DAR5 primers. PCR products were then digested with *BstXI* (NEB). Digested PCR products were separated by electrophoresis on a 6% TBE polyacrylamide gel. The gel was then dried and exposed to film.

DAR5 (28-mer forward primer):



5'-TGCTTCCGGCTCGTATGTTGGTTGGAAT-3'

FM30 (26-mer reverse primer):

5'-CTCCATTTTAGCTTCCTTAGCTCCTG-3'

### **rAAV 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 (334). The virus was subsequently purified using a rAAV virus purification kit (VIRAPUR) and the viral titer was quantitated by qPCR. A day before infection, cells were split into  $2 \times 10^5$  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  $1 \times 10^4$  was added to the other well. 2 days after infection, 1% of the cells were plated into a 10 cm culture dish without any drug selection and used to determine plating efficiency. The remaining cells were plated into 10 cm culture dishes with 2  $\mu\text{g/ml}$  puromycin. This media was replaced 5 days later with puromycin-containing media supplemented with 5  $\mu\text{g/ml}$  of 6-thioguanine (6TG) 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 (> 50 cells) at approximately 10 to 14 days later, the plates were fixed and stained with crystal violet.

## Results

### Generation of LIG3-conditional null HCT116 cell line

To generate a HCT116 cell line that is conditionally-null for LIGIII expression, an rAAV gene-targeting methodology was adopted (334), similar to the targeting strategy that had been successfully used to generate a Ku86 conditionally-null HCT116 cell line (117). The conditional targeting vector contained three LoxP sites flanking the neomycin-resistant gene (NEO) and exon 4 of LIGIII, respectively (Fig. 1A). Eight correctly targeted first round clones (LIGIII<sup>NEO/+</sup>; Fig. 1B) from 210 G418-resistant clones were identified (relative gene targeting: 3.8%). One of these clones was transiently treated with pGK-Cre to remove the NEO selection cassette (Fig. 1C). The resulting LIG3<sup>flox/+</sup> cell line was subjected to a second round of gene targeting using a regular knockout vector, where exon 4 was designed to be replaced by the NEO gene (Fig. 1C). Ten second round, correctly-targeted cell lines were identified from 711 G418-resistant clones (relative gene targeting: 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.*, LIGIII<sup>flox/NEO</sup>; Fig. 1D), which was subsequently infected with an AdCre virus to remove the NEO gene. The resulting LIG3<sup>flox/-</sup> cell line (Fig. 1E) was viable and when needed, Cre recombinase could be re-introduced to generate LIG3<sup>-/-</sup> cells (Fig. 1F).

PCR analyses were used to molecularly confirm the genetic designation of the cell lines. A 670 bp PCR product represents the wild-type LIGIII exon 4 whereas a 746 bp product should be generated when the floxed (flanked by two loxP sites) exon 4 DNA is used as a substrate. As expected, PCR of the parental wild-type LIGIII<sup>+/+</sup> cell line generated only the 670 bp band, whereas PCR of the LIG3<sup>flox/+</sup> cell line produced both the wild type (670 bp) and the floxed allele (746 bp) bands (Fig. 2A). In contrast, LIGIII conditionally null cell lines — with or without the NEO gene cassette — generated only the 746 bp PCR product corresponding to the floxed allele (Fig. 2A).

To assess the conditionality of the cell lines, LIGIII<sup>+/+</sup> and LIGIII<sup>flox/-</sup> cells were infected with increasing amount of AdCre virus, and 5 days later genomic DNA was purified and analyzed by PCR (Fig. 2B) and whole cell extracts were prepared and analyzed by immunoblotting (Fig. 2C). As the amount of Cre recombinase increased, the

PCR signal from the genomic floxed allele (764 bp) decreased until the signal was undetectable (Fig. 2B). Similarly, at the protein level, the increasing presence of Cre completely ablated detectable LIGIII protein expression (Fig. 2C). In stark contrast, Cre expression in the parental wild-type LIGIII<sup>+/+</sup> cells had no effect on formation of the exon 4-derived PCR product (670 bp) (Fig 2B) nor was any effect on LIGIII protein expression (Fig. 2C). From these experiments, we concluded that we had successfully constructed a LIG3<sup>fllox/-</sup> (*i.e.*, conditionally null) cell line.

### **The mitochondrial form of LIGIII is essential for human somatic cell viability**

To assess the essential nature of LIGIII, we initially infected the LIGIII<sup>fllox/NEO</sup> cell line with AdCre and then isolated 22 individual G418-sensitive colonies by limiting dilution. Theoretically, two possible cell lines could have been recovered: LIGIII<sup>fllox/-</sup>, and LIGIII<sup>-/-</sup>. Interestingly, all 22 recovered clones had a LIGIII<sup>fllox/-</sup> genotype and none were LIGIII<sup>-/-</sup>. This extreme asymmetry in the recovery of Cre-treated survivors suggested emphatically that the LIGIII<sup>-/-</sup> cell line was not viable (data not shown).

Because the mitochondrial form of LIGIII is essential in mice (269, 271), we tested whether this activity was conserved in human LIGIII. To this end, we complemented the LIGIII<sup>fllox/-</sup> cells with a modified LIGIII cDNA that could be expressed only in the mitochondria (mL3), and generated a stable LIG<sup>fllox/-:mL3</sup> cell line. The mL3 expression construct was made by mutating the second and third LIGIII ATGs to ATCs (Fig. 3A). The nuclear form of LIGIII is normally translated from the second ATG, and without the second ATG only the longer mitochondrial-specific version of the protein should be made (262). 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<sup>fllox/-:mL3</sup> cells with AdCre virus, we isolated by limiting dilution single cell clones, and 25 out of 43 of the isolated clones were genotypically LIG3-null (LIG3<sup>-/-:mL3</sup>) cells — a result that was opposite of what we had obtained when we tried to establish LIGIII-null cells in the absence of mL3 expression (described above). A mitochondrial exclusive expression pattern of mL3 was verified by IHC (immuno-histochemistry). Cells were first incubated with a primary antibody directed against LIGIII and then a green-fluorescent secondary antibody was used for visualization. DAPI staining was used to

visualize the nuclei. In the parental wild-type HCT116  $LIGIII^{+/+}$  cells, LIGIII protein was expressed ubiquitously (Fig. 3B). In  $LIG3^{-/-:mL3}$  cells, however, fluorescent signal was detected exclusively in the cytoplasm (Fig. 3B). From these experiments we concluded that it is solely the mitochondrial form of LIGIII that is essential for human somatic cell viability whereas the nuclear form of LIGIII is apparently dispensable for survival.

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

With a viable LIGIII-null cell line in hand (*i.e.*,  $LIG3^{-/-:mL3}$  cells) we were equipped to investigate the phenotypes resulting from the loss of nuclear LIGIII expression. Before beginning these analyses, however, we augmented our reagents with a derivative cell line that re-expressed a nuclear-specific LIGIII cDNA. This expression construct was generated by deleting the N-terminal MLS coding sequence from the wild-type cDNA and by adding a C-terminal FLAG epitope-tag to generate the nuclear-only LIGIII cDNA, nL3 (Fig. 3A). We used this construct to isolate a stable  $LIG3^{-/-:mL3:nL3}$  cell line that exhibited a strong LIGIII IHC signal from the nucleus in addition to pan-cytoplasmic staining (Fig. 3B).

### **A LIGIII deficiency causes a growth defect**

The growth rate of LIGIII-deficient cells was determined. 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 were counted on days 4 to 8 (Fig. 4).  $LIGIII^{fllox/-}$  cells showed a slight haploinsufficiency for growth, which was exacerbated in the  $LIG3^{-/-:mL3}$  cell line. From these experiments we concluded that the absence of nuclear LIGIII expression results in proliferation defects.

### **LIGIII-null cells are not sensitive to DNA damaging agents**

Given that LIGIII has been implicated in both (single-strand break) and DSB repair pathways (191, 194), we used colony-forming assays to examine the sensitivity of LIGIII-null cells to a variety of DNA damaging agents. Etoposide is a topoisomerase II inhibitor and a powerful radiomimetic that induces DNA DSBs (365). As a positive

control, a *LIGIV*<sup>-/-</sup> cell line (Chapter 2) was included. As expected, *LIGIV*<sup>-/-</sup> cells were exquisitely sensitive to etoposide, even at the lowest concentration (Fig. 5A). In contrast, *LIG3*<sup>flox/-</sup>, *LIG3*<sup>-/-:mL3</sup> and *LIG3*<sup>-/-:mL3:nL3</sup> did not show any increased sensitivity to etoposide compared to the wild-type parental cells (Fig. 5A). Similarly, the *LIGIII*-deficient cell lines were not hypersensitive to MMS (methyl methanesulfonate) (Fig. 5B) nor CPT (camptothecin) (Fig. 5C). MMS is an alkylating agent that induces SSBs and at high doses can cause DSBs (384). CPT is a topoisomerase I inhibitor that can cause DNA SSBs, DSBs and replication fork stalling. In conclusion, and somewhat surprisingly, the absence of *LIGIII* (and presumably the A-NHEJ DSB repair pathway) did not result in a frank hypersensitivity to DNA damage agents.

### **The absence of *LIGIII* does not affect the overall DNA end joining activity of human cells**

The DNA end-joining activity of *LIGIII*-null cells was measured using an extra-chromosomal reporter assay system: the pEGFP-Pem1-Ad2 construct (27, 194, 254, 343). In this assay, end joining is measured by the reconstitution of GFP expression (Fig. 6). 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. 6A). Without modification, the GFP gene is not expressed because the Ad2 exon is incorporated into GFP mRNA (Fig. 6C). 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. 6B). Productive end joining of the linear plasmid after it is transfected into the experimental cell line can be quantitated using FACS (fluorescent activated cell sorting) analysis of GFP expression (Fig. 6C). Un-digested or partially digested plasmids normally retain the Ad2 exon and therefore cannot express functional GFP protein.

When the pEGFP-Pem1-Ad2 plasmid was transfected into the parental wild-type cell line approximately 35% of the DNA was productively end joined regardless of whether it

had been digested by either *HindIII*- or *I-SceI*-digested (Fig. 6D). As expected, *LIGIV*-null cells were profoundly impaired in this DNA end-joining activity and showed only a few percent of activity in comparison to wild type. This observation confirms that *LIGIV* (and presumably C-NHEJ) is required for most of the end-joining activity detectable in human somatic cells (254). In contrast, *LIGIII*<sup>fllox/-</sup> and *LIGIII*<sup>-/-:mL3</sup> cell lines had DNA end-joining activities similar to wild-type cells (Fig. 6D). Thus, partial or complete deficiencies in *LIGIII* do not appear to affect the overall DNA end joining activity of human somatic cells.

### **Microhomology-mediated end joining is still detectable in human somatic cells lacking *LIGIII* expression**

To extend these observations, we next attempted to specifically quantitate microhomology-mediated end-joining. To this end, a reporter substrate, pDVG94, that is biased towards detecting microhomology-mediated end-joining events (23, 254, 385) was used *in vivo* to measure the end-joining activity of *LIGIII*-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. 7A). Repair of this substrate by C-NHEJ generally generates a product that retains at least some of either repeat, whereas microhomology-mediated end-joining (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 then 24 hr later the DNA was recovered from the cells and repaired junctions were amplified by PCR with radiolabeled primers (Fig. 7A). 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 end-joining (23, 254, 385). As expected, and as has been previously reported (254), wild-type cells generated predominately the 180 bp product, consistent with most of the end-joining in human somatic cells resulting from C-NHEJ. Similarly, and again as expected (Chapter 2), a *LIGIV*-null cell line showed highly elevated levels of the 120 bp product indicative of a reliance on A-NHEJ. Surprisingly, in either *LIGIII*<sup>fllox/-</sup> or two independent *LIGIII*<sup>-/-:mL3</sup> cell lines the amount of

the 120 bp product was unchanged in comparison to wild-type cells (Fig. 7B) indicating that LIGIII-deficient cells are still able to carry out microhomology-mediated end-joining.

### **A LIGIII deficiency does not affect the overall rAAV gene-targeting rate**

AAV infections in humans result in no known pathogenesis and this feature (amongst others) has made rAAV gene targeting technology one of the more promising candidates for therapeutic use (327, 329, 331). Nonetheless, random viral integration events are clearly not desired and thus increasing the correct gene targeting frequency is one of the most sought-after advances for rAAV-mediated gene targeting technology. Previous studies from our laboratory had indicated that neither HR nor C-NHEJ were likely responsible for rAAV random integrations (276). In addition, sequencing results from other laboratories had indicated that many, if not all, rAAV random integration events were mediated by microhomology usage. Together, these observations led to the hypothesis that rAAV random integrations are mediated by A-NHEJ, and that perhaps by disrupting A-NHEJ (*e.g.*, by functionally inactivating LIGIII) the correct rAAV-mediated gene-targeting rate could be improved.

To experimentally test this hypothesis, we performed gene targeting in a LIGIII-null cell line at the HPRT (hypoxanthine phosphoribosyltransferase) locus using an rAAV gene targeting vector designed to disrupt exon 3 of HPRT. HPRT is an X-chromosome linked gene and the enzyme encoded by HPRT plays a central role in the generation of purine nucleotides through the purine salvage pathway (386). Cells expressing a wild-type HPRT gene are 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 only a single round of gene targeting, 6-TG selection could be used to isolate the correctly targeted clones. Interestingly, LIGIII-null cells showed a 2-fold increase in the frequency of correct targeting (Fig. 8B). Unexpectedly (and unfortunately) the frequency of random integration events was not reduced in LIGIII-null cells, but was actually enhanced (Fig. 8A). Overall, there was no statistical difference in the targeting rate between wild-type and LIGIII-null cell lines (Fig. 8C). These data

demonstrate that while LIGIII is a general suppressor of rAAV integrations, it does not preferentially affect random versus correct targeting events.



## Discussion

### LIGIII has an essential mitochondrial function

In this study we generated a viable human somatic cell line that lacks the expression of nuclear LIGIII. Somewhat paradoxically, we nonetheless conclude that LIGIII is an essential gene. This conclusion is based on the presumed requirement for isoforms of LIGIII in mitochondrial maintenance. Thus, no live null cells could be isolated after an AdCre infection of the LIGIII<sup>fllox/NEO</sup> cell line. In contrast, after complementing the LIG3<sup>fllox/-</sup> cell line with a LIGIII cDNA that encoded a mitochondrial-specific isoform of LIGIII (mL3), we easily isolated LIGIII-null (LIGIII<sup>-/-:mL3</sup>) cells. Therefore, we conclude that in human somatic cells, LIGIII is dispensable in the nucleus but essential in mitochondria. This conclusion is consistent with a recent report describing a similar conditionally-null knockout strategy of LIGIII in the mouse (269).

What the essential activity of LIGIII in the mitochondria may be, is currently unclear. One obvious function would be a requirement for LIGIII in Okazaki fragment maturation during mitochondrial DNA replication. LIGI mediates Okazaki fragment maturation in the nucleus (192, 387, 388), but since LIGI lacks a MLS and it is not detected in mitochondria, it is clear that some other ligase must perform this function in mitochondria. Similarly, since LIGIV also lacks a MLS, is non-essential (Chapter 2) and appears to be involved exclusively in C-NHEJ it is also a poor candidate. In contrast, the mitochondrial-specific isoform of LIGIII should be able to mediate the ligation of Okazaki fragments, which are very similar to the intermediates that LIGIII is known to ligate together during BER (247, 389). Curiously, there does not appear to be a specific requirement for LIGIII in the mitochondria — *i.e.*, its essential nature is due perforce to the fact that it is the only ligase expressed in the mitochondria. In a very revealing experiment, a chimeric expression construct consisting of the MLS of LIGIII fused to the coding region of LIGI functionally rescued the mitochondrial defects of LIGIII-null mice (269). Thus, while there is a requirement for ligase activity in the mitochondria of mammals, there does not appear to be a requirement for a specific ligase.

Other explanations are also possible. Thus, for example, mitochondria are the sites of oxidative phosphorylation and this process generates a lot of ROS (reactive oxygen species) and/or free radicals (390, 391). ROS, in turn, are very potent generators of DNA

DSBs (392, 393), which, if they accrued to a high enough level, would certainly have deleterious consequences on the mtDNA (mitochondrial DNA) in terms of gene expression and DNA replication. It is therefore fairly easy to envision that a dedicated DSB repair system must be in place to keep mtDNA intact. Again, it is unclear whether LIGIII would assume this role by default since it is the only ligase localized to mitochondria and/or whether ROS-induced damage generates structures that can only be ligated by LIGIII.

### **rAAV random integrations are not mediated by LIGIII/A-NHEJ**

The mechanism of gene targeting is intimately linked to DNA DSB repair pathways. Correct targeting events require HR and the random integrations of the targeting vector are presumably facilitated by DNA DSB end-joining pathways (*i.e.*, C-NHEJ, A-NHEJ or both). In published work from our laboratory, reductions in Ku did not significantly impact the frequency of random integrations of rAAV vectors when they were used for gene targeting (276). Similar observations have been made for LIGIV- (chapter 2), XLF- (F. Fattah *et al.*, manuscript in preparation) and XRCC4- (B. Ruis and E. A. Hendrickson, unpublished data) deficient cells suggesting very strongly that C-NHEJ is not required for this process. Thus, we anticipated that impairing the A-NHEJ DSB repair pathway should greatly reduce rAAV random integrations and increase the overall gene targeting rate. Surprisingly, no such effect was observed (Fig. 8B) and in fact, the random rAAV integration frequency was actually somewhat elevated, demonstrating that A-NHEJ is normally a suppressor of these events. This result is perplexing. One possibility is that there is genetic redundancy such that in C-NHEJ-defective cell lines, random integration events are carried out by A-NHEJ and in A-NHEJ-defective cells they are carried out by C-NHEJ. We are currently trying to construct a LIGIII-:LIGIV-doubly-null cell line (*i.e.*, defective for both A- and C-NHEJ pathways, respectively) to experimentally address this possibility (although see below for issues related to this construction). A potentially more interesting possibility is if there exists yet another — hitherto — unidentified end-joining pathway. This hypothesis is actually supported by our results analyzing the DNA repair capacity of LIGIII-null cells (see below). If this model is correct than the LIGIII-

:LIGIV-doubly-null cell line (assuming it is viable) may still be proficient for random rAAV integration events.

Finally, it is worth noting that previous work has shown that defects in any of the C-NHEJ genes do not improve the frequency of rAAV-mediated correct targeting events. The only exceptions to this are cells defective in either Ku subunit, where significantly elevated levels of rAAV-mediated gene targeting occur (276). These experimental results have been interpreted to mean that Ku is the master regulator of DSB repair pathway choice and can essentially determine the ultimate chromosomal location of an incoming piece of dsDNA. This model would predict that the status of LIGIII is irrelevant to this process. In contrast, we observed that the correct gene targeting frequency mediated by rAAV was elevated ~3-fold in LIGIII-null cells (Fig. 8A). This increase in gene targeting is not nearly as large as that seen in Ku-defective cells (276, 277, 369), which would be consistent with Ku being the major regulator of the gene targeting process. Nonetheless, this increase also suggests strongly that Ku can't be the sole regulator and it warrants more investigation into how LIGIII mechanistically impacts this process.

### **The existence of a LIGIII-independent MMEJ pathway**

We utilized a reporter construct, pDVG94, to measure MMEJ activity (Fig. 7). In order to generate a 120 bp *BstXI*-dependent restriction product from this plasmid a unique repair event is required: both ends of the blunt-ended plasmid must be resected and 6 nt long complementary strands must be precisely annealed and ligated without the loss or addition of a single nucleotide (Fig. 7A). This repair product can be detected only at very low levels (a few percent of the total) in wild-type cells (Fig. 7B) and has been widely interpreted as being produced by A-NHEJ (23, 254). When cells are mutated genes required for C-NHEJ, this 120 bp fragment becomes virtually the sole repair product (Fig. 7B; Chapter 2; (254)). All of these observations led to the prediction that the ablation of A-NHEJ should block the formation of this repair product. Our data show unequivocally that this does not happen. Thus, the appearance of the 120 bp product was completely unaffected (and perhaps even slightly elevated) by the absence of nuclear LIGIII (Fig. 7B). This was, once again, an unexpected and perplexing result. One

possibility is that our cell line is leaky and does in fact still express some LIGIII in the nucleus. While we cannot 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 LIGIII, but a downstream ATG as well to ensure that not even truncated forms of the protein would be expressed (Fig. 3A). In addition, the mitochondrial-exclusive expression pattern of the LIGIII<sup>mL3</sup> construct was confirmed by IHC (Fig. 3B). Finally, it is hard to reconcile the slow growth phenotype of LIGIII-null cells (Fig. 4) with the possibility that enough LIGIII<sup>mL3</sup> was expressed to enable A-NHEJ. Still it should be pointed out that if LIGIII<sup>mL3</sup> protein did leak into nucleus, even at undetectable levels, this could likely have had functional relevance since it has been demonstrated that even low levels of LIGIII are sufficient for effective A-NHEJ (394). Finally, it should be noted that this explanation is also relevant to the lack of an effect that was observed by the loss-of-function of LIGIII on gene targeting described earlier.

An alternative attractive possibility is that there is another end-joining pathway, discrete from C-NHEJ and A-NHEJ. 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 LIGIII (Fig. 7B). A more complicated explanation could be if the C-NHEJ and A-NHEJ pathways mechanistically overlap. Thus, these results could be explained if the C-NHEJ pathway has the capacity to produce low levels of the 120 bp product irrespective of A-NHEJ. This model predicts that A-NHEJ is essentially inactive in the presence of functional C-NHEJ — as has indeed been argued by numerous investigators (23-25, 27, 159, 244, 249, 347, 395). In this scenario, the inactivation of LIGIII would have no effect on the production of the low levels of the 120 bp product as long as C-NHEJ is active; a result that is consistent with our data (Fig. 7B). Again, one genetic approach to answering this question would be to try and construct a LIGIII<sup>-</sup>:LIGIV-doubly mutant cell line and measure the residual end-joining activity.

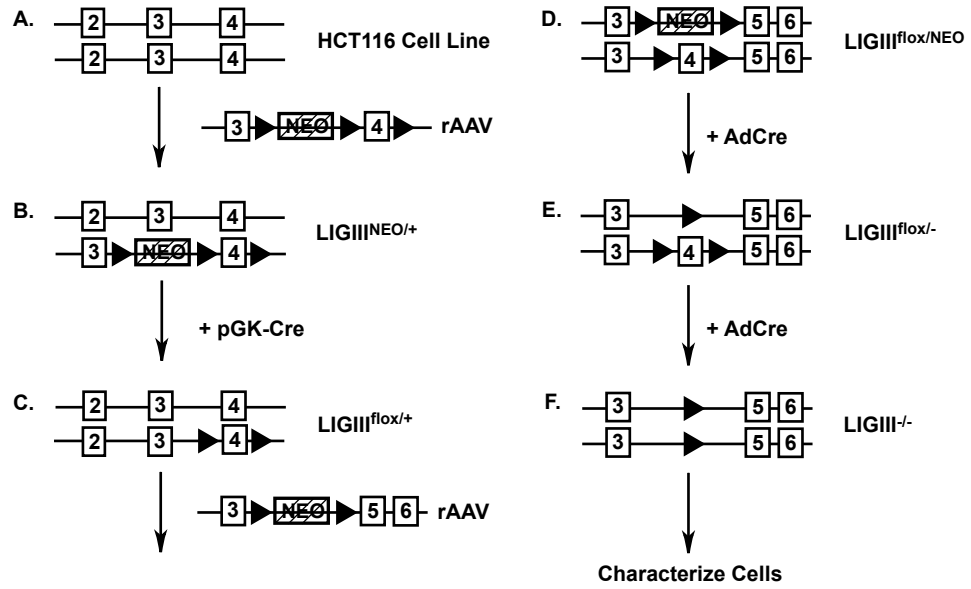
### **A deficiency in both C-NHEJ and A-NHEJ induces synthetic lethality**

One of the reasons that made it hard to identify any A-NHEJ defects in our LIGIII-null cells was because the majority of repair events were mediated by C-NHEJ. As noted

above, we have begun genetic experiments to try and address these issues. These complicated doubly mutant genetic experiments, do require, however, a lot of time. In an attempt to obtain a more expeditious, albeit less rigorous, answer we tried to suppress C-NHEJ by transient siRNA knockdowns. Specifically, we knocked-down the expression of DNA-PK complex subunits in the LIGIII-null background. Unfortunately, however, knockdown of either Ku70 or DNA-PK<sub>cs</sub> in LIGIII-null cells induced a serious proliferation defect that was ultimately accompanied by high levels of apoptotic cells (data not shown). The high cell mortality made it technically impossible to characterize the cells any further. Importantly, the efficient knockdown of Ku70 to a level that was less than 5% of wild-type (254, 276), and the complete ablation of DNA-PK<sub>cs</sub> expression by targeted knockout (363) did not by themselves cause significant apoptosis in contrast to what was observed in the LIGIII-null background. Although preliminary, these experiments at least suggest that the deficiency of both C-NHEJ and A-NHEJ in human somatic cells may not be tolerated. Future genetic experiments are directed at addressing this issue.

**Figure 1.** Targeting strategy and primer sets used for the screening. (A) A cartoon of part of the LIGIII genomic locus. The rectangles represent exons and lines in between are introns. The conditional knockout vector has three loxP sites (triangle) that flanks a NEO cassette and exon 4. (B) G418-resistant clones were selected and among them, correctly targeted clones were identified by PCR. LIGIII<sup>NEO/+</sup> clones were treated with pGK-Cre to remove the NEO gene prior to the second round of targeting. (C) A G418-sensitive LIGIII<sup>fllox/+</sup> clone was targeted with an exon 4 knockout vector. (D) A correctly targeted clone, LIGIII<sup>fllox/NEO</sup>, was treated with AdCre to derive a G418-sensitive LIGIII<sup>fllox/-</sup> clone, depicted in (E). (F) AdCre treatment of LIGIII<sup>fllox/-</sup> clone results in LIGIII<sup>-/-</sup> cells, which were used for characterization.

Figure 1. Scheme for functional inactivation of the human LIGIII locus.



**Figure 2.** AdCre infection removes both LIGIII DNA and protein in LIGIII<sup>fllox/-</sup> cell. (A) PCR test of LIGIII<sup>fllox/-</sup> cell. We used the primer set that gives different sized PCR products from loxP-floxed allele (746 bp) and wild-type allele (670 bp). With the same primer set, NEO gene will give over 3 kb band and loxP-only allele will give smaller than 200 bp product, but they are not shown in this figure. (B) LoxP-floxed allele was disappeared after AdCre infection. With increasing amount of AdCre virus, 746 bp of loxP-floxed allele is gone in LIGIII<sup>fllox/-</sup> clone. (C) Western blot shows the loss of LIGIII protein. Protein samples were parallel to the DNA samples in (B). Increasing amount of AdCre virus is correlated with decreasing amount of LIGIII protein.



Figure 2 (A) PCR test of a DNA LIGIII conditional null cell line

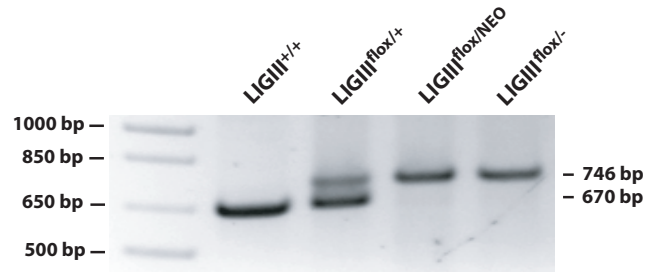
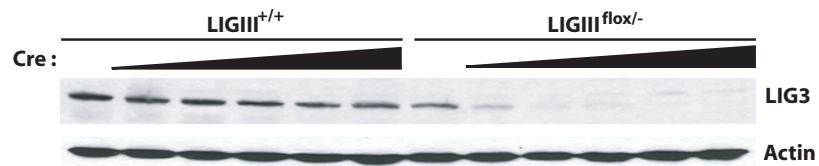


Figure 2 (B) PCR characterization of the DNA LIGIII conditional null clone after AdCre infection

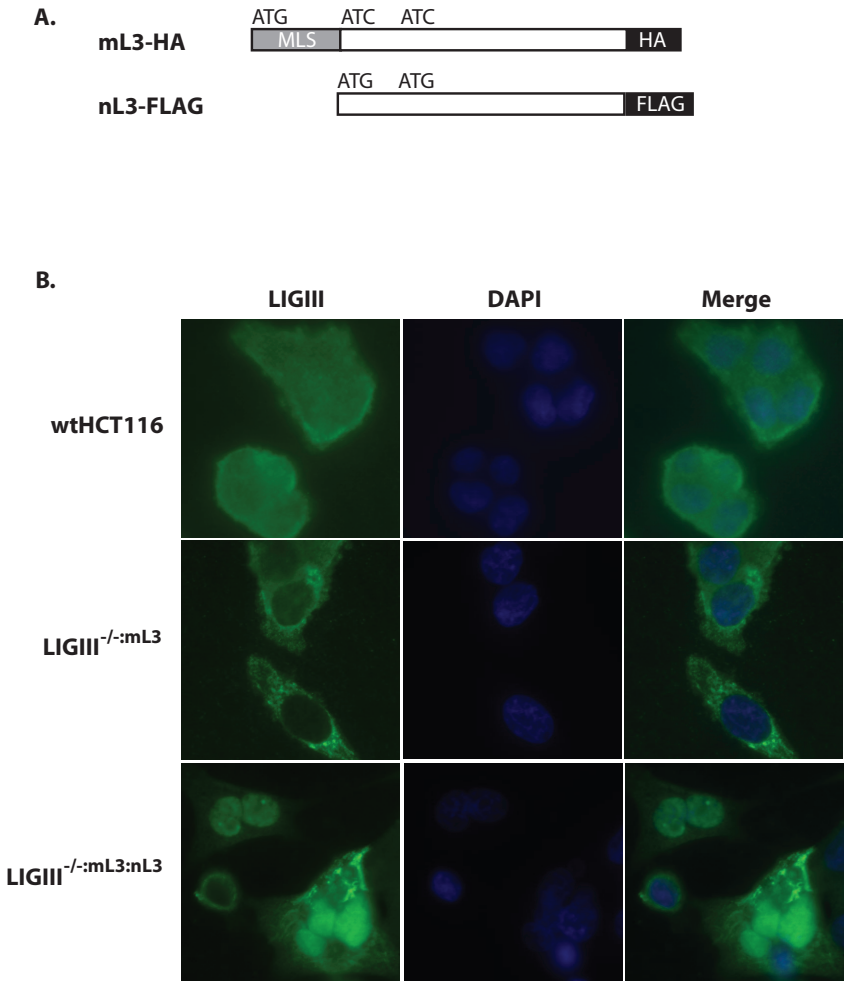


Figure 2 (C) Western blot analysis of the DNA LIGIII conditional null clone after AdCre infection



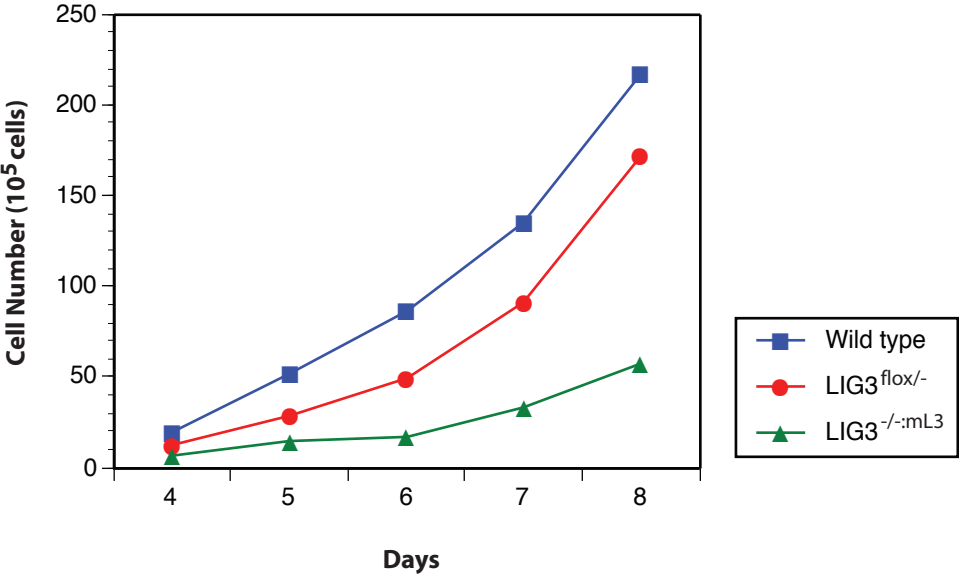
**Figure 3.** Complementation with mitochondria-only LIGIII cDNA rescues the lethality of LIGIII<sup>-/-</sup> cells. (A) Constructs used for complementation. To make a mitochondrial-only LIGIII (mL3), the second and third ATGs were mutated to ATC. For the nuclear-only LIGIII (nL3) construction, the N-terminal MLS sequence was deleted. mL3 and nL3 were modified with C-terminal HA- and FLAG- epitope tags, respectively. (B) Nuclear and mitochondrial localization of LIGIII complemented clones. In wild-type HCT116, LIGIII is expressed ubiquitously. In contrast, LIGIII<sup>-/-:mL3</sup> cells showed a mitochondrial-exclusive expression pattern. In LIGIII<sup>-/-:mL3:nL3</sup> cells, mL3 is detected in the cytoplasm and nL3 is over-expressed in the nucleus.

**Figure 3. Complementation with a mitochondrial-exclusive form of LIGIII rescues the cell lethality**



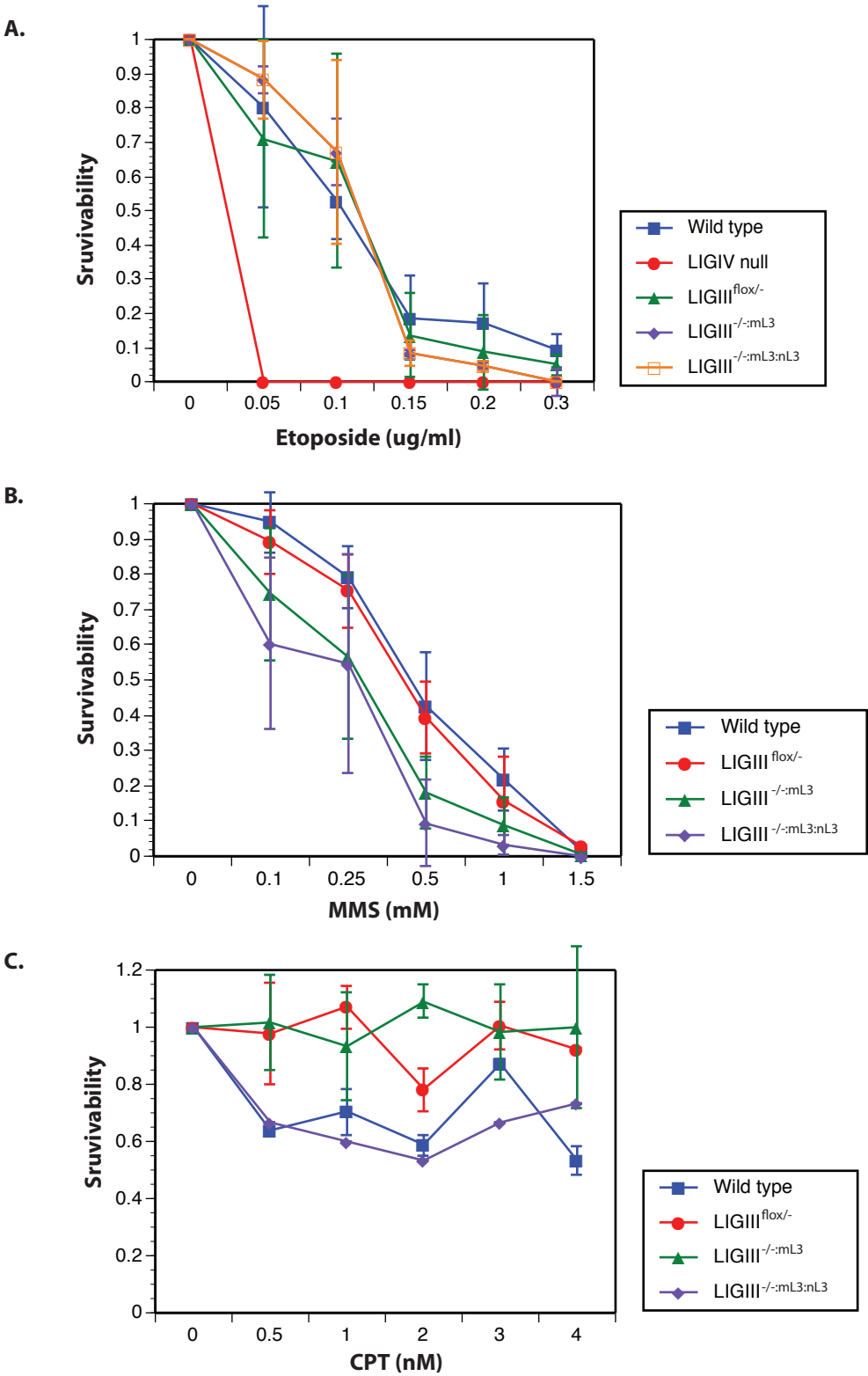
**Figure 4.** Growth defects of human somatic cells with reduced or no LIGIII. Three thousand cells were plated on day 0 and their growth was assessed by counting trypan blue excluding cells at the indicated days. The average of two independent experiments, each done in triplicate, is shown.

Figure 4. LIGIII deficiency causes growth retardation



**Figure 5.** LIGIII deficient cells are not sensitive to DNA damaging agents. (A) Etoposide sensitivity test. Three hundred cells were plated a day before etoposide treatment. Clone #35 is LIGIII<sup>-/-:mL3</sup> cell, and clone #34 is LIGIII<sup>-/-:mL3:nL3</sup> cell. Survival was normalized by setting the value from no etoposide treatment as 1. As a positive control, a LIGIV<sup>-/-</sup> cell line, which is known to be extremely sensitive to etoposide, was used. The plotted values are the average of three independent experiments. (B and C) MMS and CPT sensitivity tests, respectively, were 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, and CPT-containing media for 24 hr.

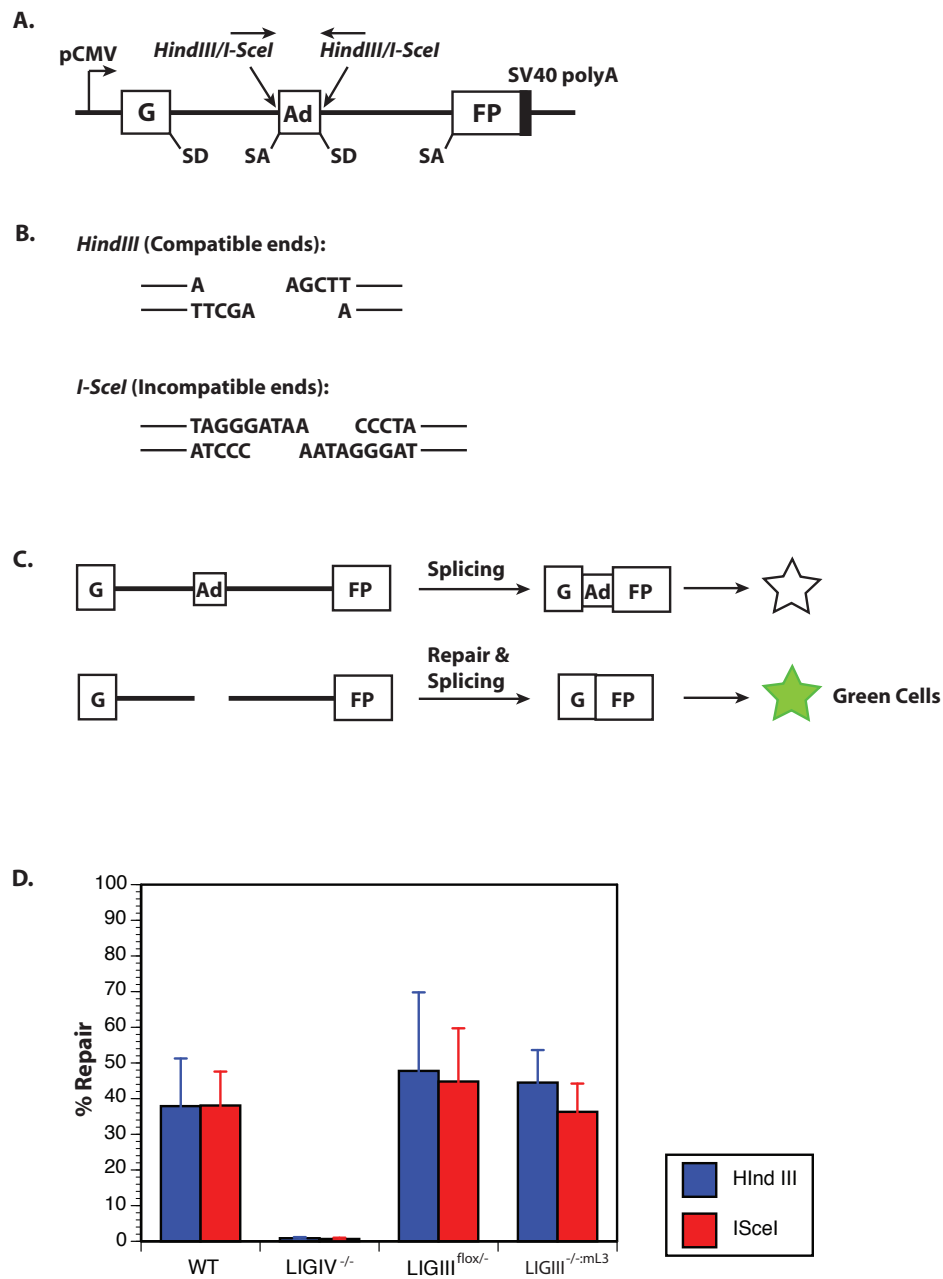
**Figure 5. LIGIII-null cells do not show sensitivity to DNA damaging agents**



**Figure 6.** A LIGIII deficiency does not affect end-joining activity. (A) A cartoon of the pEGFP-Pem1-Ad2 repair substrate used for analysis of NHEJ 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 shown. (B) *HindIII* digestion generates compatible ends with 4 nt overhangs, whereas *I-SceI* digestion produces ends, that require some processing before they can be rejoined. (C) Due to the presence of the Ad-exon into the middle of the Pem1 intron, the Ad exon is efficiently spliced into the middle of the GFP ORF, inactivating the GFP activity and thus making the starting substrate GFP negative. Both sides of the Ad exon have *HindIII/I-SceI* restriction sites. Cleavage with either of these endonucleases 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 LIGIII deficiency on end joining.

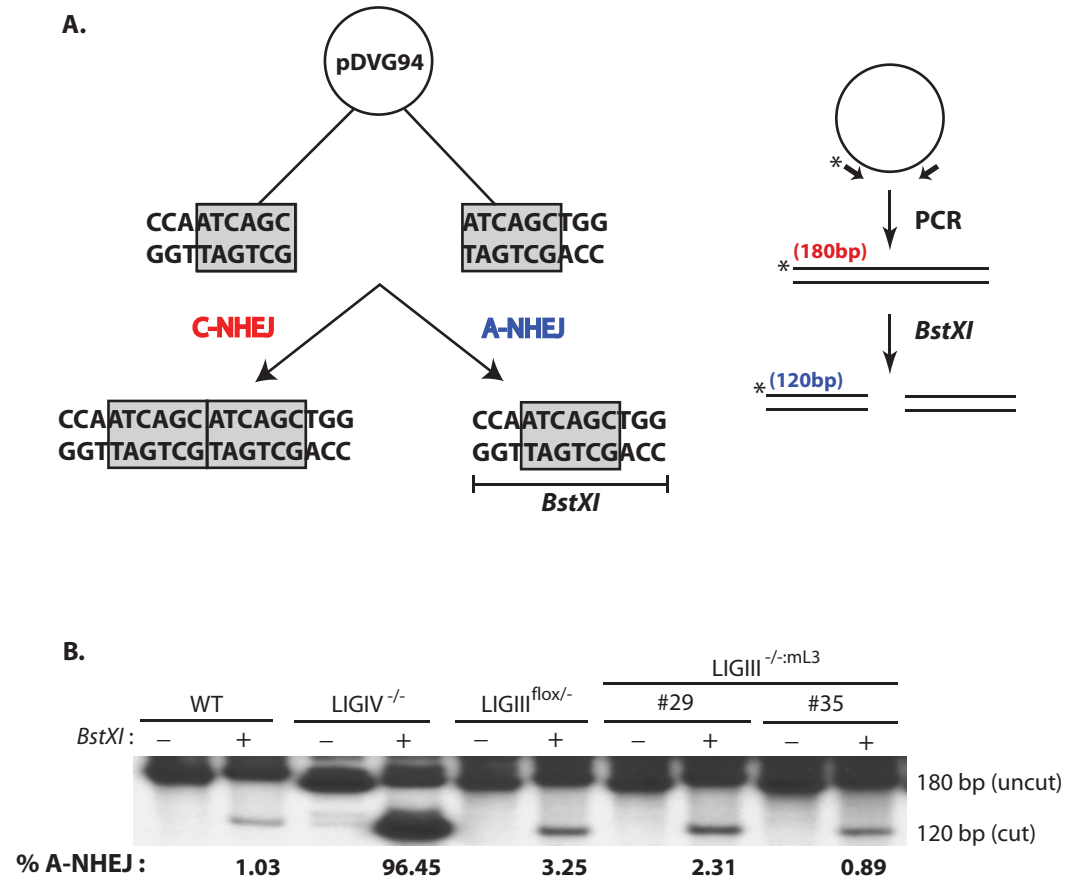


Figure 6. A *LIGIII* deficiency has no effect on the overall DNA end-joining ability of human somatic cells



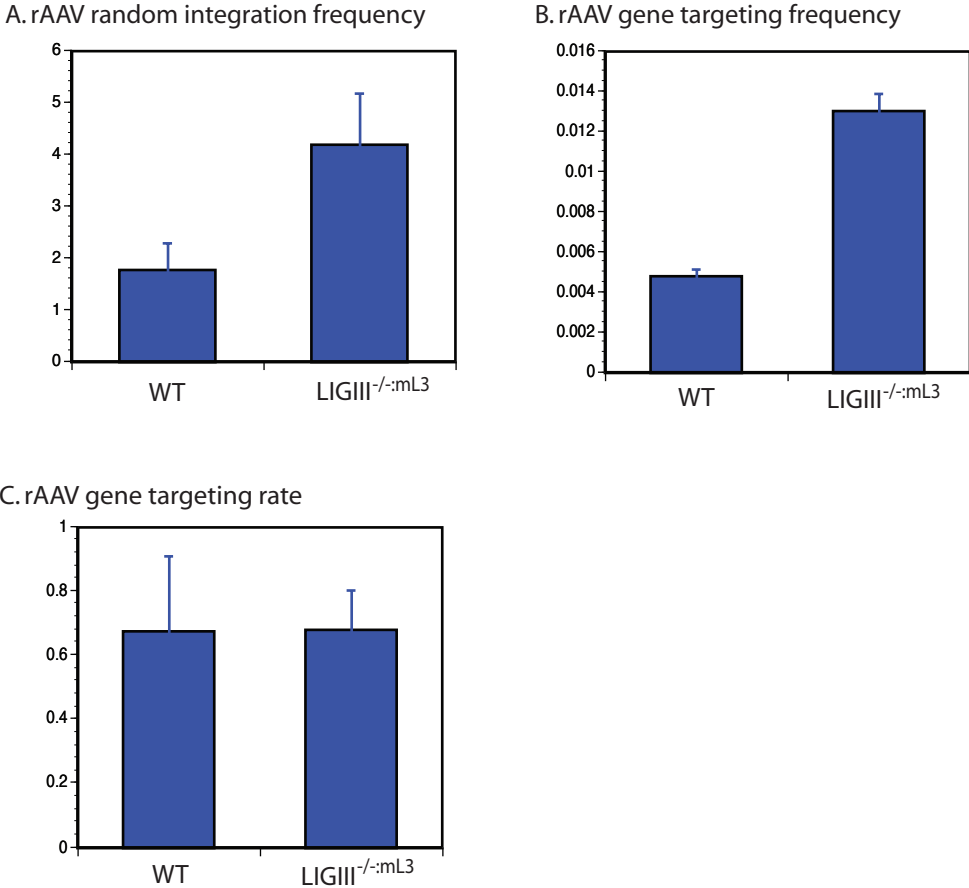
**Figure 7.** LIGIII-null cells have normal microhomology-mediated end-joining activity. (A) After *EcoRV* and *AfeI* restriction enzyme digestion, the reporter substrate, pDVG94, becomes a blunt-ended linear plasmid with 6 bp direct repeats at both ends. Repair via C-NHEJ generally retains at least part of either repeats, whereas A-NHEJ generates only a single repeat and a novel restriction enzyme site which can be cleaved by *BstXI*. Repaired junctions were amplified by PCR using radiolabeled primers and the 180 bp PCR product was subjected to *BstXI* digestion. The 180 bp uncut product represents repair via C-NHEJ whereas the 120 bp digested product represents A-NHEJ mediated repair. (B) Both LIGIII-het and LIGIII-null cells have similar microhomology end-joining activity to the wild-type control.

**Figure 7. LIGIII-null cells still retain microhomology-mediated repair activity**



**Figure 8.** A LIGIII deficiency does not affect the rAAV-mediated gene targeting rate at the HPRT locus. (A) rAAV random integration frequency. The random integration frequency was determined as the number of Puro-resistant clones normalized by plating efficiency. (B) rAAV gene targeting frequency. The gene targeting frequency was determined as the number of both Puro- and 6TG-resistant clones normalized by plating efficiency. (C) The rAAV gene targeting rate represents the ratio of the correct targeting events versus the total viral integration events. LIGIII-null cells show both higher random integration and gene targeting frequencies, but overall, the LIGIII-null cells did not show any difference in gene targeting rate compared to wild-type cell.

**Figure 8. A** *LIGIII* deficiency has no effect on the overall rAAV gene targeting rate



## **CHAPTER IV**

**A deficiency in Rad54B rescues the telomere defects and leads to the prolonged survival of Ku86-null cells**

All eukaryotes maintain their DNA in the form of linear chromosomes. Linear chromosomes — or more precisely, the ends of linear chromosomes — however, have attendant problems including issues about how they are replicated and how they can be distinguished from interchromosomal DSBs (double-strand breaks). To solve these problems, all eukaryotic cells have evolved a repetitive DNA structure, termed a “telomere” that resides at the ends of each chromosome. In human cells, a telomere consists of 4 to 14 kb of a hexameric (TTAGGG)<sub>n</sub> repeat that ends with 200 to 300 nt of a 3'-G-rich, single-stranded overhang. This overhang can invade the neighboring duplex telomeric DNA to form a lariat-like, D-loop structure that is referred to as a t-loop. A set of dedicated telomere-specific binding proteins — collectively called “shelterin” — bind to the t-loop. The combination of the unusual structure and the proteinaceous cap formed by shelterin appear to protect the ends of chromosomes from random nuclease activity and from recognition by DNA DSB repair machineries. Because one of the main roles of telomeres is to sequester chromosomal ends away from DNA DSB repair factors it came as a large surprise when several damage signaling and DNA repair proteins were shown not only to be physically associated with telomeres, but to be required for their proper maintenance. Among these, Ku86 was one of the most unexpected components, because it is the essential DNA end-binding factor for C-NHEJ (classic non-homologous end joining), the main DNA DSB repair pathway. Nonetheless, the loss-of-function of Ku86 results in cell lethality not because of a defect in C-NHEJ (although that certainly happens) but due to the telomere defects, including telomere loss, sister chromatid fusions, and the loss of the t-loop in the form of a circle (*i.e.*, “t-circle” formation). The formation of t-circles has been experimentally observed in several other situations and it has been generally inferred that this happens because of the aberrant action of HR (homologous recombination) on t-loops. To experimentally test this hypothesis, we generated a Ku86:Rad54B (RAD54B plays an essential role in HR) doubly-null human cell line. Ku86:Rad54B doubly-null cells displayed a greatly attenuated lethality and significantly reduced amounts of t-circle formation. Thus, our results demonstrate that in human somatic cells the t-circle formation mediated by the loss-of-function of Ku86 is largely a RAD54B- (and by extrapolation, HR-) dependent process.

## Introduction

Most cells encounter almost constant challenges to their genomic stability in the form of DNA damage and were it not for equally robust DNA repair activities, all cells would rapidly succumb to the accumulated damage (396). One of the most difficult challenges a cell faces is how to deal with a DNA DSB. Importantly, the genomic instability caused by DNA DSBs results in insertions, deletions, and — probably most pathologically — translocations, which are known to predispose cells to a cancerous fate (273). DNA DSBs can arise from different sources (2). There are endogenous sources, such as V(D)J recombination, a site-specific recombination reaction required for the development of the immune system (167), stalled replication forks, and ROS (reactive oxygen species) that is the byproduct of oxidative phosphorylation in mitochondria. Cells can also suffer DSBs due to exogenous sources. The most clinically relevant of these is exposure to IR (ionizing radiation), which is often used as a therapeutic modality in the treatment of cancer (397). Although the absolute number of DSBs per cell on a daily basis is likely to be fairly low (398), their potential to generate pathological outcomes means that each and every DSB needs to be properly repaired as studies in yeast have demonstrated that even a single unrepaired DSB can be a lethal event (338).

Not surprisingly, all living organisms have evolved robust mechanisms to efficiently repair DSBs. In mammals there are least three genetically distinct DNA DSB repair pathways: HR (homologous recombination), C-NHEJ (classic non-homologous end joining) and A-NHEJ (alternative NHEJ) (340). HR requires significant end resection and it uses extensive sequence homology, which results in essentially error-free repair products (10). In contrast, C-NHEJ does not necessarily require any end resection and very frequently results in error-prone repair with associated deletions and/or insertions (18). A-NHEJ appears mechanistically to be somewhat of a hybrid between the other two pathways; requiring some, but not necessarily significant amounts of, resection and the use of small, but not extensive, blocks of homology, which result in a repaired product in which a deletion has always occurred (286). In human cells, the C-NHEJ pathway appears to be the highly preferred pathway of DNA DSB repair (2). C-NHEJ is highly researched and the mechanism is well understood. There are seven major C-NHEJ factors and almost all of these genes are evolutionarily conserved from bacteria to



humans. Somewhat confusingly, however, loss-of-function mutations in any of these factors can lead to different phenotypes in different organisms. The best example of this concerns mutations of either Ku subunit, Ku70 or Ku86. In all organisms — with one glaring exception — in which either one of these genes has been functionally inactivated, the resulting cells or animals are viable, albeit with varying hypersensitivities to DNA damaging agents and immune deficits. For example, Ku70- and Ku86-null mice, while presenting with IR-hypersensitivity and immune deficits are nonetheless viable (160, 161). In striking contrast, neither null, nor even heterozygous, human Ku70- or Ku86-mutant patients has ever been reported. Moreover, inactivation of either Ku70 or Ku86 in human somatic cell lines resulted in death (124). This species-specific difference was very puzzling until our laboratory demonstrated that the lethality of the human Ku86 deficiency was caused by telomere dysfunction (117).

Telomeres, the repetitive DNA sequence found at the ends of all linear chromosomes seemed to have evolved with the emergence of eukaryotes, where linear chromosome replaced the circular chromosome of prokaryotes. The transition from circular to linear chromosomes solved some old problems concerning chromosome segregation, but simultaneously created some new problems related to chromosome synthesis and stability (105). Because the DNA replication of the lagging chromosomal strand occurs by Okazaki fragment synthesis, the logistics of this process requires that a chromosome end will be shortened each cell cycle, resulting, ultimately, in the loss of genetic information. The other problem associated with linear chromosome ends is end-protection. Linear chromosome ends are structurally very similar to DNA DSBs and the “repair” of chromosome ends will cause chromosome fusion and, ultimately, other types of genetic instability. Telomeres solve both of these problems. Human telomeres consist of 4 to 14 kb of hexameric (TTAGGG)<sub>n</sub> repeats that terminate with a 200 to 300 nt single-stranded, G-rich overhang (298). The G-rich overhang performs two quite important functions. First, it is the substrate for a dedicated reverse transcriptase complex, termed telomerase, than can extend chromosome ends when they shorten (399). Second, it invades and hybridizes to the duplex part of telomere forming a classic, lariat-shaped D-loop, termed a t-loop. The somewhat bizarre t-loop structure in turn acts as a binding site for a six-protein complex, called shelterin. The unusual structure of telomeres and the

proteinaceous cap formed by the shelterin complex essentially makes the chromosomal ends invisible to the cellular DNA DSB repair pathways and thus imparts stability to them. Very paradoxically, some of the same DNA repair and damage signaling proteins that can fuse short or uncapped telomeres are required for normal telomere and t-loop maintenance (302). This is best exemplified by Ku86, one half of the potent DNA end-binding factor, Ku, used by C-NHEJ (112, 113, 119, 142, 371). Short and/or uncapped telomeres will be fused by C-NHEJ in a Ku-dependent fashion, resulting in either cell death or cellular transformation. Remarkably, however, human Ku is also required for the maintenance of normal telomeres. Human cells lacking Ku86 die due to telomere dysfunction, including massive telomere loss, sister chromatid fusions, and the conversion of the t-looped structure into a circles (*aka* a “t-circle”) (117). The formation of t-circular DNA has been reported in a variety of conditions in which telomere dysfunction has been induced (117, 400) and the general consensus is that t-circle formation results from the aberrant action of HR on the t-loop. Since Ku deficient cells are known to result in an upregulation of HR (276, 369, 401-405), one hypothesis is that the elevated HR levels in Ku-deficient cells leads indirectly to the t-circle formation.

To experimentally investigate this hypothesis, we devised a scheme to analyze telomere function in HR and Ku86 doubly-deficient cell. To this end, we deleted the Rad54B (radiation sensitive 54B) gene to inactivate HR. Rad54B is a member of the Rad52 epistasis group, which encodes the core factors of HR (67, 81). Rad54B was first identified as a novel member of the SNF2/SWI2 ATP-dependent chromatin remodeling family and a homolog of Rad54, which is a DNA-dependent ATPase that functions as a chromatin remodeling factor during HR (81). Rad54B has been previously inactivated in a human cell line and the resulting viable cell line did not have any appreciable growth defects or sensitivities to DNA damaging agents. Importantly, however, HR activity appeared to be completely abrogated based on gene targeting experiments where the targeting frequency dropped from 7 to 15% to undetectable (83). Thus, we functionally inactivated the Rad54B gene in the existing Ku86 conditionally-null HCT116 cell line (117). Impressively, the deficiency in Rad54B suppressed the formation of t-circles and greatly extended the lifespan of cells converted to the Ku86-null phenotype. Based on these observations, we conclude that RAD54B (and by logical extension, HR) does

indeed mediate the telomere dysfunction caused by Ku86 loss-of-function mutations in human cells.

## Materials and Methods

### Rad54B targeting vector construction and screening

The pAAV-Neo-Rad54B targeting vector was generated as described (334). The 661 bp of DNA needed for the left homology arm was made by PCR using the E3-LARM-NotI-F1 and E3-LARM-SacII-R1 primers. The 1620 bp of right homology arm DNA was made by PCR using the E3-RARM-KpnI-F1 and E3-RARM-NotI-R3 primers. Primer sets #1 and #2 were subsequently used for screening for correctly targeted clones. Primer set #1 consisted of Rad54B-E3-CreS-F1 and Rad54B-E3-CreS-R1. Primer set #2 consisted of Rad54B-E3-LARM-SC-F4 and NeoR2.

E3-LARM-NotI-F1:

5'-ACATAAGCGGCCGCTTTAAGTATTGATTTTAGTATTGAGAAATTTAAC-3'

E3-LARM-SacII-R1:

5'-GGCGGCCCGCGGCTAAAAGAAACAAATATATATTTAAATCAGAACTC-3'

E3-RARM-KpnI-F1:

5'-CCGGTACCGACTGCTTTTTATTGATAAGGTTTATGCTTGACC-3'

E3-RARM-NotI-R3:

5'-ACATAAGCGGCCGCGGTGATGGGGAAAATGACATATGTTATTTAACTGG-3'

Rad54B-E3-CreS-F1:

5'-GAGTTCTGATTTAAATATATATTTGTTTCTTTTAG-3'

Rad54B-E3-CreS-R1:

5'-CAAGCATAAACCTTATCAATAAAAAGC-3'

Rad54B-E3-LARM-SC-F4:

5'-CCAACATAGTGAGATTACCATTATCTCACC-3'

NeoR2:

5'-AAAGCGCCTCCCCTACCCGGTAGGGCG-3'

### Cre recombination of the Ku86 conditionally-null cell lines

Ku86-null HCT116 cells were generated when needed as described with a slight modification (117). Ku86<sup>fllox/-</sup> cells were plated in 6-well plates at 5 X 10<sup>4</sup> cells/well and

16 to 18 hr later,  $5 \times 10^8$  viral particles were added in 2 ml media. After infection (24 hr), the virus-containing media was replaced with 2 ml of fresh media.

### **Use of an inducible Cre system**

A PiggyBac transposon system (406) was used with slight modification. Cells were subcultured into 6-well plates a day before transfection. A vector expressing the PiggyBac-transposase (2  $\mu$ g) and another vector containing the PiggyBac-CreERt2-transposon lacking a GFP (green fluorescent protein) marker (2  $\mu$ g) were transfected using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. After transfection (2 days), the cells were subclutured into a 96-well plate using 2  $\mu$ g/ml puromycin in the media for selection. Clones stably expressing CreER were subsequently identified by immunoblotting using an antibody directed against Cre (data not shown). For the CreEr induction, a 1 mM 4-OHT (tamoxifen) stock solution, which was dissolved in ethanol, was diluted 1000 fold.

### **Immunoblotting**

Whole cell extract was prepared using RIPA buffer (Sigma) and proteins were subjected to electrophoresis on a 4 to 20% gradient gel (Bio-Rad). Ku86 (SantaCruz) and actin (SantaCruz) antibodies were used at 1:500 and 1:250 dilutions, respectively. A Cre-antibody (Abcam) was used at 1  $\mu$ g/ml and a Rad54B antibody (gift of R. Kanaar) was used at a 1:1000 dilution. UV-irradiated HeLa cell lysate (40  $\mu$ g; SantaCruz) was used as a positive control.

### **Growth rate**

On day 0,  $5 \times 10^4$  cells were plated in 6 cm culture dishes. Starting at day 1, cells were, where required, incubated in 5 nM of 4-OHT-containing media. At later time points, cells were harvested and counted using a Countess cell counter (Invitrogen).

### **Neutral-neutral, 2D-gel electrophoresis**

Neutral-neutral, 2D-gel electrophoresis was performed as described (117).

## **T-FISH**

T-FISH (telomere-FISH) was performed with metaphase-arrested cells using a protein-nucleic acid telomere-specific probe {Cy3 conjugated to  $(T_2AG_3)_3$ }, according to the protocol provided by the manufacturer (DAKO) and previous studies (115, 117).

## Results

### Construction of a Rad54B gene targeting vector

The HCT116 cell line has three copies of the Rad54B gene because a portion of one chromosome 8 — which is where the Rad54B gene resides — is duplicated (83). This required three rounds of gene targeting and screening (Sup. Fig. 1). Exon 3 was chosen for deletion because its removal produces an out-of-frame mutation. The homology arms for this vector were asymmetric, with only 661 bp in the left homology arm whereas the right homology arm was considerably longer at 1620 bp (Fig. 1A). The length of the homology arms, which are normally of equal length (276, 334, 363), were adjusted in this way to accommodate extensive repetitive sequences present upstream of exon 3. In general, we try to avoid including repetitive sequences in the homology arms when at all possible, but at the same time we try to maximize the total length of homology arms, because a longer total homology in the arms increases the efficiency of rAAV gene targeting (407).

For screening, we used two different sets (#1 and #2) of primers (Fig. 1A). Primer set #1 anneals to a region outside of exon 3 and generates different-sized PCR products for the wild-type allele and the loxP-only allele, which are 237 bp and 146 bp, respectively. Correctly targeted clones were subsequently treated with the Cre recombinase to excise the NEO selection cassette, so that the same targeting vector could be used for the following round of gene targeting. A loxP-only allele was generated by this Cre recombination reaction and was detected by the generation of a 146 bp product using primer set #1. Primer set #2 uses a forward primer that anneals upstream of exon 3 and a reverse primer that anneals to the NEO selection cassette. This generates a diagnostic 768 bp PCR product only from correctly targeted clones. PCR analyses following a third round of gene targeting using primer set #2 (Fig. 1B) and #1 (Fig. 1C), respectively confirmed the isolation of the desired RAD54B-null clones. In these experiments, genomic DNA generated from Rad54B<sup>+Neo/-</sup> cells was used as a positive control for PCR. Clone #63 corresponded to a retargeted clone because the 237 bp PCR product produced from the wild-type allele was still present (Fig. 1C). In contrast, clones #84, #91, and #103 were Rad54B<sup>-/-</sup> clones, because they were correctly targeted (Fig. 1B) and they no longer had any exon 3 sequences (Fig. 1C).

Ku86 and Rad54B protein expression was then verified by immunoblot analysis. HeLa cell extract was used as a positive control (Fig. 2A). Parental (wtHCT116) cell extract was also used a positive control, which was important since this genetic background contained a cross-reacting, non-specific band that migrated near the position of authentic RAD54B (Fig. 2A). Rad54B<sup>-/-</sup> HCT116 cell extract was used as a negative control. Consistent with the PCR analyses ((Fig. 1C) clone #63 (a re-targeted clone) still expressed RAD54B protein (Fig. 2A) whereas the putative Ku86<sup>flox/-</sup>:RAD54B-null clones #9, #91 and #103 did not express detectable Rad54B protein (Fig. 2A). Ku86 protein levels were also assessed in these clones along with the parental (WT) and Ku86<sup>flox/-</sup> HCT116 control cell lines. All of the cell lines were infected with AdCre and the levels of Ku86 expression were determined by immunoblot analysis at day 5. Ku86 protein expression in wild-type cells did not change even after AdCre infection (Fig. 2B). However, all the Ku86<sup>flox/-</sup>:Rad54B clones and the Ku86<sup>flox/-</sup> clone expressed almost no Ku86 protein after AdCre infection, confirming that their Ku86 alleles were conditionally null (Fig. 2B). From these analyses, we concluded that we had successfully generated viable Ku86<sup>flox/-</sup>: Rad54B<sup>-/-</sup> cell lines.

### **A Rad54B deficiency in a LIGIV-null background causes synthetic lethality**

Initially we tried to functionally inactivate the Rad54B gene in LIG4<sup>-/-</sup> as well as Ku86<sup>flox/-</sup> cells. The single LIGIV<sup>-/-</sup> mutant cell line is viable ((207); Chapter 2) and it is essentially devoid of C-NHEJ activity ((254); Chapter 2). The single Rad54B<sup>-/-</sup> mutant cell line is also viable (see above) and completely defective in HR (83). Because each single mutant was viable, we hoped that the LIGIV<sup>-/-</sup>:Rad54B<sup>-/-</sup> doubly mutant cell line would also survive and serve as “A-NHEJ-only” system. Unfortunately, the simultaneous absence of LIGIV and Rad54B in HCT116 cells caused a synthetic lethality. We reached this conclusion based upon gene targeting data. Thus, the first and second rounds of RAD54B targeting in the LIGIV<sup>-/-</sup> cell line targeting occurred with a 8.8% and 13.2% correct targeting efficiency, respectively (Table 1A). An asymmetry in the targeting frequency was already noted, however, during the second round of targeting as 8 out of 12 clones were re-targeted at the previously inactivated allele (Table 1A). This asymmetry was greatly exacerbated during the third round of targeting. The overall



targeting rate dropped to only 4.7 %, and 10 out of 11 correctly targeted clones were shown to be re-targeting events (Table 1B). Moreover, even though one clone was Rad54B<sup>-/-</sup> based on PCR analysis, it proliferated very poorly and then stopped growing altogether after a month, which made it impossible to characterize further (Table 1A).

These results stand in stark contrast to the data we obtained for inactivating RAD54B in the Ku86<sup>flox/-</sup> background. In these experiments, the gene targeting frequency remained constant at ~5% and there was no re-targeting bias observed in the second and third rounds (Table 1B). Together, these data strongly suggest that a combined LIGIV and Rad54B deficiency triggers a synthetic lethality and imply that the A-NHEJ pathway is not sufficient to sustain human cell survival. On the other hand, cells expressing only 50% the parental levels of Ku86 are decidedly still viable even with a RAD54B deficiency. Thus, human cells can apparently survive on A-NHEJ, as long as they have some C-NHEJ function.

### **Introduction of a 4-OHT-inducible Cre-recombinase system**

Visual observation of Ku86<sup>flox/-</sup>:Rad54B<sup>-/-</sup> cells following an AdCre infection suggested that many of the cells were still alive at times when all of the Ku86<sup>flox/-</sup> cells were dead or dying (data not shown). This suggested that the Rad54B deficiency might rescue the lethality of the Ku86 loss-of-function. Attempts to isolate viable Ku86<sup>-/-</sup>:Rad54B<sup>-/-</sup> cells, however, 2 to 3 weeks after AdCre infection yielded only cells with a Ku86<sup>flox/-</sup>:Rad54B<sup>-/-</sup> genotype, resulting — presumably — from cells where Cre recombination did not occur. The inability to isolate stable Ku86<sup>-/-</sup>:Rad54B<sup>-/-</sup> cells suggested that a Rad54B deficiency simply delayed the death of Ku86<sup>-/-</sup> cells. Using AdCre methodologies to pursue this hypothesis, however, was fraught with several technical difficulties. First, AdCre infections turn out empirically to be “tricky”, where the precise amount of AdCre is required to yield low toxicity and effective recombination. Small errors on either side of the optimal concentration yield either very high toxicity or too low of a conversion rate of the flox to the null allele, respectively. Second, even though adenovirus can be quite an effective system at optimal doses, it is never 100% efficient. Empirically, in AdCre-treated cultures, the unrecombined cells eventually overgrow the null cells and repopulate the culture. To solve these technical

problems, we introduced a 4-hydroxyl tamoxifen (4-OHT)-inducible Cre-recombinase into our cells using a PiggyBac transposon delivery system. Two vectors are required for this system. The first vector is capable of transiently expressing the PiggyBac transposase *in trans* (Fig. 3A). The other vector contains the Cre-recombinase linked to the estrogen receptor at its C-terminus (CreER) and a puromycin selection marker, which are flanked by 5'- and 3'-PiggyBac terminal repeats (Fig. 3A). These two constructs were co-transfected into the relevant cell lines and puromycin-resistant clones were isolated and characterized for CreER expression by Western blotting (data not shown). Clones (*e.g.*, #1 and #19, below) expressing CreER were designated as “CreER”.

The induction efficiency of this system was first determined.  $Ku86^{lox/-}$  and  $Ku86^{lox/-:CreER}$  cell lines were treated with 4-OHT or the vehicle, EtOH. Clones #1 and #19 were two independent clones of  $Ku86^{lox/-:CreER}$ , both of which expressed robust levels of CreER protein. The viability of  $Ku86^{lox/-}$  cells was affected by neither EtOH nor 4-OHT (Fig. 3B). In contrast, the  $Ku86^{lox/-:CreER}$  cells were completely killed when Cre expression was induced by 4-OHT, but not by EtOH (Fig. 3B). These encouraging results demonstrated that the system is specifically responsive to 4-OHT and not leaky. Moreover, unlike the adenovirus infections, which by themselves can be quite toxic, the 4-OHT treatment itself did not have any detectable toxicity (Fig. 3B). Next, we checked the toxicity of the Cre recombinase expression. We treated  $WT^{CreER}$  cell and  $Ku86^{lox/-:CreER}$  cells with increasing concentrations of 4-OHT. After 7 days of 4-OHT treatment, the  $Ku86^{lox/-:CreER}$  cells had all succumbed whereas no defects in the proliferation of the  $WT^{CreER}$  cells was detected, indicating that overexpression of Cre recombinase in the nucleus is not inherently toxic to cells.

### **The absence of Rad54B increases the survivability of Ku86-null cells**

With the CreER system in hand, we set out to determine how a Rad54B deficiency affects the lethality of Ku86-null cells. Relevant clones stably expressing CreER were treated with 4-OHT for 12 days. The viability of  $WT^{CreER}$  cells was essentially unaffected by a 4-OHT treatment (Fig. 3D). In contrast, 4-OHT exposure killed almost all of the  $Ku86^{lox/-:CreER}$  cells (Fig. 3D). Intriguingly, the viability of  $Ku86^{lox/-:CreER}:Rad54B^{-/-}$  clones, while poor, was nonetheless clearly higher than  $Ku86^{lox/-:CreER}$  treated cells (Fig.

3D). The absence of Ku86 protein in the  $Ku86^{\text{flox/-};\text{CreER}}$  and  $Ku86^{\text{flox/-};\text{CreER}};Rad54B^{-/-}$  cells after 4-OHT treatment was confirmed by immunoblot analyses (Fig. 3E). These results demonstrate that a deficiency in Rad54B partially rescues the lethality induced by the absence of Ku86. This conclusion was verified by measuring the growth rate (Fig. 4). Although the rescue was slight compared to wild-type cells (Fig. 4A), it was significantly better than the survival of the Ku86-null cells (Fig. 4B).

#### **A Rad54B deficiency suppresses the t-circle formation that occurs in Ku86-null cells**

Telomeric DNA, when aberrantly acted upon by HR, can be lost in the form of t-circles. To test whether the enhanced survival associated with the absence of RAD54B (and by extrapolation, HR; (83)) in Ku86-null cells correlated with diminished t-circle formation we performed neutral/neutral 2D gel electrophoresis using  $Ku86^{-/-}Rad54B^{-/-}$  and control cells. As a positive control, asynchronously growing WI-38 VA13 cells were analyzed. VA13 is an ALT (alternative lengthening of telomeres) cell line that is known to constitutively express t-circular DNA (117, 306, 408). Indeed, a retarded arc on the 2D gels was observed (Fig. 5A) as had been reported (117). Moreover, genomic DNA purified from  $Ku86^{-/-}$  cells 5 days after 4-OHT treatment showed a well-defined t-circular arc. In contrast, the  $Ku86^{-/-}Rad54B^{-/-}$  cell line did not contain t-circles (Fig. 5C). These experiments demonstrated that t-circle formation in  $Ku86^{-/-}$  cells requires RAD54B and, likely, the entire HR pathway.

## Discussion

### A Rad54B loss-of-function partially rescues the lethality of Ku86-null cells

In this study, we generated a Ku86<sup>fllox/-</sup>:Rad54B<sup>-/-</sup> cell line and found that the lethal phenotype of Ku86-null human cells can be partially rescued in the absence of RAD54B/HR activity. Human Ku86-null HCT116 cells die because of severe telomere dysfunction, including the production of t-circular DNA (117). In contrast, human Rad54B-null HCT116 cells are viable (with normal proliferation), but they contain no functional HR activity, based on the observation that they are incapable of performing gene targeting (83). We observed that when we combined the absence of RAD54B with the Ku86-null genotype that this delayed the death of the Ku86-null cells (Fig. 3). In addition to the extended survival, Ku86<sup>-/-</sup>:Rad54B<sup>-/-</sup> cells showed significantly reduced amounts of t-circles compared to the single mutant Ku86<sup>-/-</sup> cells (Fig. 5). These results are consistent with models where 1) t-circles are formed by HR and 2) where Ku86 acts as a potent suppressor of HR at telomeres.

The original description of aberrant t-circle formation was observed in a study where TRF2 $\Delta$ B, which lacked its N-terminal basic (“B”) domain, was over-expressed in mouse cells (306). The authors of this study concluded that TRF2 could act as a suppressor of t-circle formation, presumably by suppressing the action of HR on t-loops. The mechanism of this suppression, however, was not clear. Our laboratory subsequently showed that the t-circle formation could be phenocopied by the inactivation of Ku86 (117). We came to a similar interpretation as the previous group; but again, the mechanism by which Ku86 might mediate the HR suppression was not clear. Our current work and additional work from other laboratories clarifies some, albeit not all, of these points. First, it is now known that Ku and TRF2 proteins directly interact (112, 141). A simple interpretation of this data is that TRF2 may simply act as the sequence-specific binding factor that is needed to physically localize Ku to telomeres. This model is consistent with the fact that TRF2 $\Delta$ B is known to retain all of its DNA binding activity. The regions of TRF2 and Ku that are needed for interaction are not yet defined, but we propose that Ku either interacts directly with the “B” domain or that the “B” is needed for TRF2’s interaction with Ku. In other words, our model predicts that the t-circle formation observed in TRF2 $\Delta$ B cells is due to the absence of Ku at the telomeres and not

directly due to the expression of TRF2 $\Delta$ B. We regard Ku as a much likelier candidate than TRF2 for the actual protein that suppresses HR, because this is a function that Ku is known to possess at non-telomeric sequences. Thus, Ku-reduced cells have highly elevated levels of correct gene targeting at a bevy of different chromosomal loci (276). If Ku can suppress HR at intragenic locations it seems likely that it can also do so at the end of chromosomes. Needless to say, our work has yet to elucidate precisely how Ku mediates this suppression of HR, but we favor a model in which Ku physically interacts with, and inhibits, a key HR factor. That t-circle formation requires HR activity now seems irrefutable; at least in some species or cell types. Earlier researchers transiently knocked down the expression of XRCC3 or NBS1, factors implicated in HR (408), and demonstrated that t-circle formation induced by TRF2 $\Delta$ B overexpression (306) or in ALT cells (400) was concomitantly reduced. Our “cleaner” loss-of-function genetic experiments lend great support to this earlier work and strongly implicate HR as the culprit in t-circle formation. Nonetheless, it needs to be noted that there are reports of HR-independent t-circle formation. Thus, there are reports in WRN (Werner’s Syndrome) patient cells and in some plants that t-circle formation can proceed in a XRCC3 or NBS1-independent fashion (409, 410). These processes clearly need to be investigated in more depth, but at face value one might conclude that there exist two pathways for t-circle formation: one HR-dependent and one HR-independent. If this is the case, then our data strongly suggest that the t-circle formation induced by the loss of Ku86 expression utilizes the HR-dependent pathway. Viewed in this light, it will be interesting to make double mutants with Ku86<sup>flox/-</sup> and either XRCC3 or NBS1. Our prediction would be that removal of either of these factors would also suppress the lethality of Ku86-null cell lines.

### **The mechanism of the prolonged survival of Ku86<sup>-/-</sup>:Rad54B<sup>-/-</sup> cells**

The loss of Ku86 function results in t-circle formation and eventual cell death (117). The inactivation of RAD54B in this genetic background ablates t-circle formation (Fig. 5) and prolongs the cell’s survival (Fig. 3). While provocative, this observation nonetheless begs the obvious question. That is, doubly mutant cells do not form t-circles and although they have an extended lifespan, they ultimately succumb. Thus, if the cells

are not dying due to t-circle formation, they must be dying due to some other defect. Some of our most recent preliminary data suggests that the doubly mutant cells may still be dying due to telomere dysfunction; just not t-circle formation. Experiments to characterize the G-strand overhang indicate that the G-overhang is greatly hyperextended in the doubly mutant cell line (data not shown). This observation is consistent with an earlier published report from our laboratory (115) demonstrating that Ku86<sup>+/-</sup> cell lines have elongated G-overhangs. These data suggest that Ku may actually have two functions at telomeres: suppressing HR and t-circle formation and suppressing the action of nucleases and the resulting G-overhang hyperresection. In cells proficient for HR, the loss of suppression on HR in Ku mutants predominates and that results in a deadly increase in t-circle formation. In cells defective for HR (*i.e.*, RAD54B-null cells), however, the loss of suppression on end resection predominates and the cells die due to hyperextended G-overhangs. We are currently trying to use technologies that will allow us to quantitate single-stranded telomeric DNA. In the future, these experiments could also be addressed genetically by looking for enhanced survival in cells defective in the nuclease thought to be involved in telomeric end resection such as ExoI or Dna2 (411).

#### **A deficiency in both C-NHEJ and HR is not compatible with cell survival**

The lethal phenotype of Ku86<sup>-/-</sup> cells is delayed, but not completely rescued in the absence of Rad54B. Similarly, we also demonstrated that while LIGIV-null and RAD54B-null single mutant cell lines were viable, the double LIGIV-null:RAD54B-null cell line died (Table 1A). Thus, although cells are capable of surviving without HR or C-NHEJ, they are apparently not capable of surviving without both pathways. Another way to look at this, is that A-NHEJ by itself is apparently not sufficient for cell survival. Given the mutagenic nature of A-NHEJ, this should perhaps not be such a great surprise. Nonetheless, it will be very interesting to determine if HR (*e.g.*, a LIGIII-null:LIGIV-null cell line) or C-NHEJ (*e.g.*, a LIGIII-null:RAD54B-null cell line) by themselves are capable of sustaining human somatic cells.

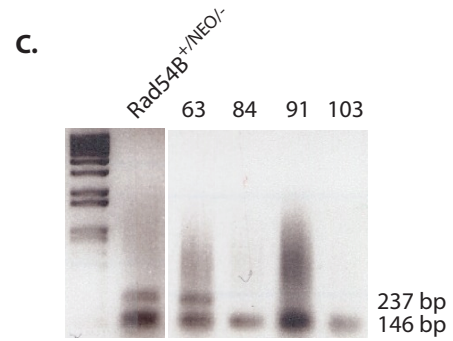
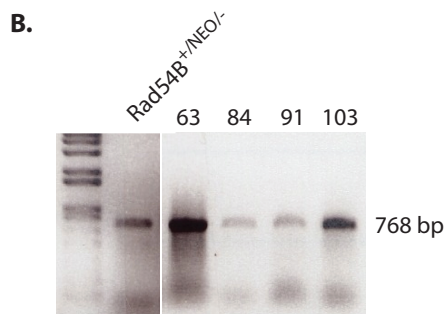
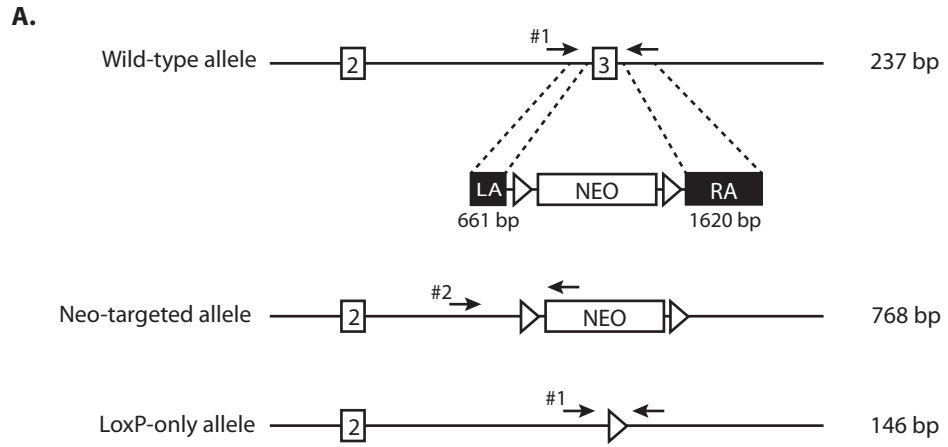
#### **The phenotype of Ku86-deficiency is species-specific**

The functional inactivation of genes in mice often fails to phenocopy what actually transpires in human cells, and telomere biology is perhaps one of the best examples of this. Ku86-null mice are completely defective for V(D)J recombination and show growth defects, but they are still alive (161, 412). In contrast, LIGIV-null mice are early embryonic lethal (209, 210, 356). Surprisingly, the embryonic lethality of LIGIV-deficient mice can be rescued by the loss of Ku86 expression (413). Ku70 and LIGIV doubly-null mice also have a less severe phenotype than LIGIV-null singly mutant mice (160). These studies are at least consistent with Ku playing a role in suppressing HR (276, 369, 401-405) and A-NHEJ (24, 159, 254). Presumably, the loss of Ku elevates the expression of HR and A-NHEJ enough to compensate somewhat for the loss of C-NHEJ due to the absence of LIGIV. In contrast to all of this work, however, Ku70 (276) and Ku86 (115, 117, 124) are essential for human somatic cell viability whereas LIG4-null human cells are viable. Moreover, the lethality of Ku86-deficient human cell comes from telomere defect, which is not rescued by a LIGIV deletion (data not shown). Ku86<sup>-/-</sup>:LIG4<sup>-/-</sup> cells die with a cell death phenotype that is indistinguishable from that of Ku86<sup>-/-</sup> cells (data not shown). These opposing phenotypes demonstrate that care should be taken in extrapolating results obtained in model systems to humans.

**Figure 1.** The Rad54B targeting strategy. (A) To generate a Rad54B knockout, exon 3 was replaced by a NEO cassette using rAAV gene targeting methods. The targeting vector had 661 bp of a left homology arm (LA) and 1620 bp of a right homology arm (RA). Two primer sets, #1 and #2, were used to identify the wild-type, targeted and loxP-only alleles. (B) The PCR products generated by primer set #2. Genomic DNA from a Rad54B<sup>+/NEO/-</sup> clone was used as a positive control for PCR. Clones #63, #84, #91, #103 are from the third round of gene targeting. (C) The PCR products generated by primer set #1. Clones #84, #91 and #103 produced only the 146 bp band, demonstrating the absence of any remaining exon 3 (237 bp) sequences.

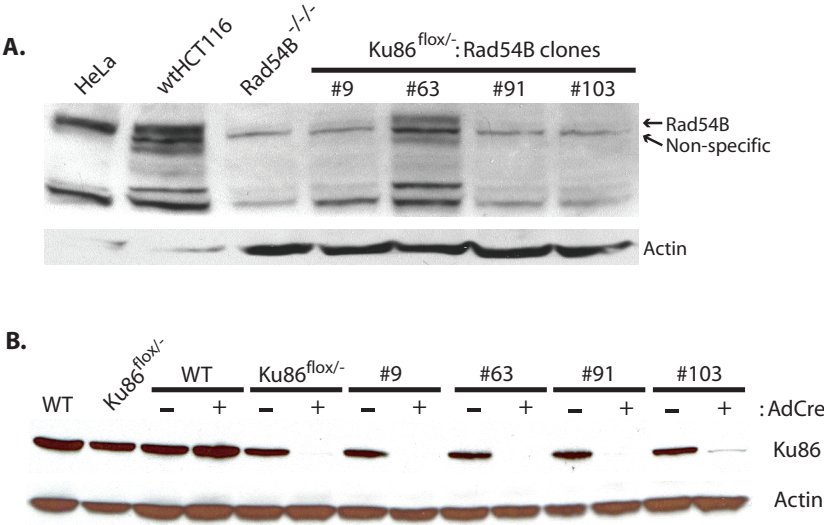


**Figure 1. The Rad54B targeting strategy**



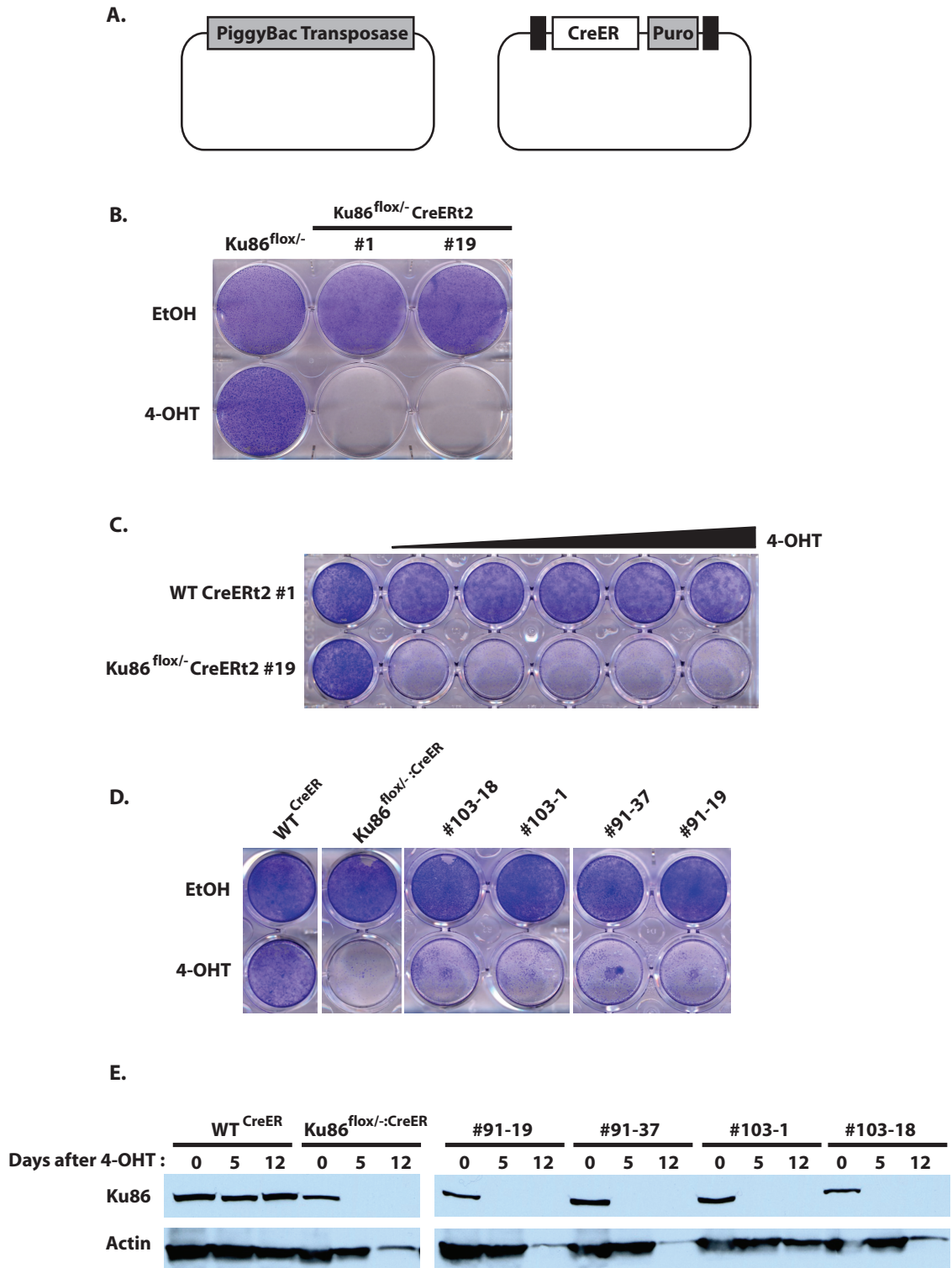
**Figure 2.** The production of Ku86<sup>-/-</sup>:Rad54B<sup>-/-</sup> cells was confirmed by immunoblotting. (A) The loss of Rad54B protein. Whole cell extracts were prepared from HeLa, wtHCT116, Rad54B<sup>-/-</sup> HCT116, and Ku86<sup>flox/-</sup>:Rad54B<sup>+/-</sup> third round targeted clones. Clones #9, #91 and #103 are Rad54B<sup>-/-</sup>. #63 is a retargeted clone, which still has one conditional Rad54B allele. Actin was used as a loading control. (B) Loss of Ku86 protein after AdCre virus infection. Ku86<sup>flox/-</sup>:Rad54B<sup>-/-</sup> cells ( $5 \times 10^4$ ) were plated a day before they were either left uninfected (-) or infected (+) with AdCre at the MOI of  $1 \times 10^4$ . After infection (5 days), cells were harvested and whole cell extracts were prepared for immunoblot analysis. Untreated wtHCT116 and Ku86<sup>flox/-</sup> cell lysates were included as controls.

**Figure 2. Western blot analysis confirms the construction of a Ku86:Rad54B doubly-null cell line**



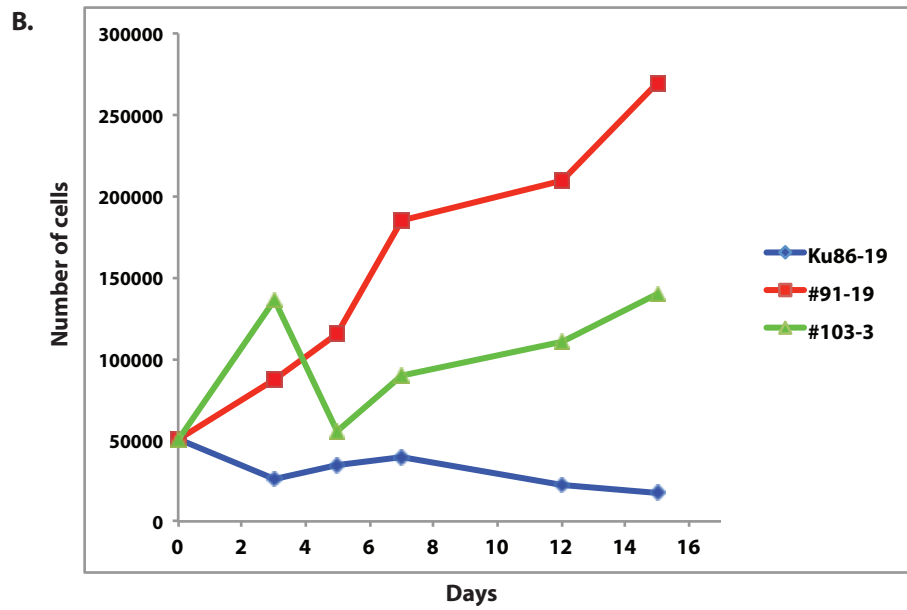
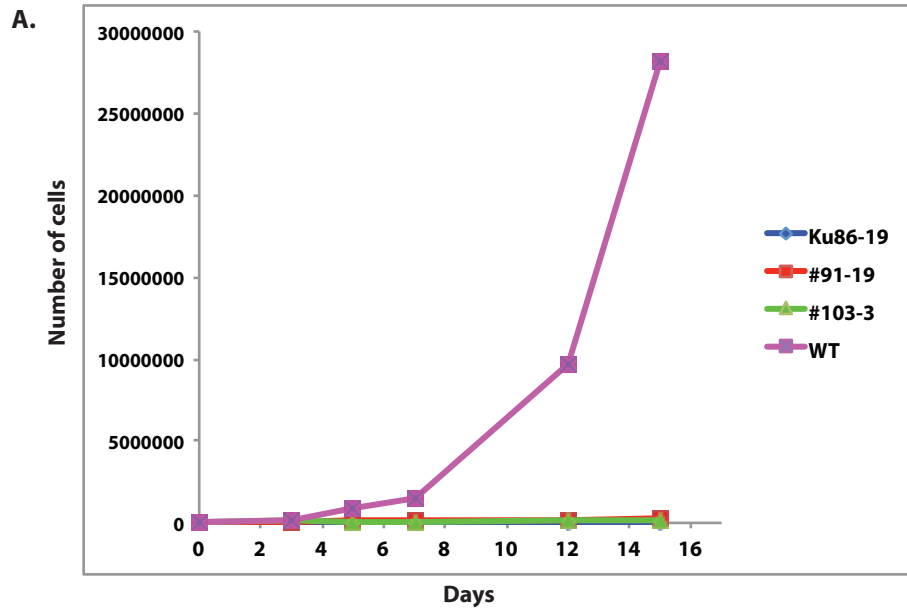
**Figure 3.** A 4-OHT-inducible CreER system was established in WT (wild-type),  $Ku86^{lox/-}$  and  $Ku86^{lox/-};Rad54B^{-/-}$  clones. (A) Cells were co-transfected with two constructs. One construct transiently expresses the PiggyBac transposase. In the other construct, the coding sequence of the cre recombinase fused with the estrogen receptor (CreER) and the puromycin-resistance gene are flanked by PiggyBac terminal repeats (black boxes). (B) The 4-OHT-inducible system efficiently causes the excision of  $Ku86$  and results in cell death.  $Ku86^{lox/-;CreER}$  cells were treated with 4-OHT or vehicle (EtOH) for 2 weeks. Cells treated with 4-OHT died whereas those treated with EtOH did not. (C)  $WT^{CreER}$  cell and  $Ku86^{lox/-;CreER}$  cells were treated with increased amounts of 4-OHT for 7 days. Only the  $Ku86^{lox/-}$  cells died. (D) Delayed death of  $Ku86^{-/-};Rad54B^{-/-}$  cells. All the indicated cells express CreER and they were treated with 4-OHT or EtOH for 12 days. Clones #103-1, #103-18 and #91-19, #91-37 are sub-clones of #103 and #91 cells, respectively. (E)  $Ku86$  protein is not expressed in 4-OHT-treated  $Ku86^{lox/-}$  cells. The indicated cell lines were treated with 5 nM of 4-OHT on day 0 and the  $Ku86$  protein level was checked on days 0, 5 and 12. 15  $\mu$ g of whole cell extracts were used for this immunoblot analysis. On day 5, there was no detectable  $Ku86$  protein in any of the clones except  $WT^{CreER}$ . On day 12, most of  $Ku86$ -deficient cells were dead, so 15  $\mu$ g of whole cell extracts could not be collected.

Figure 3. A 4-OHT-inducible CreERT2 system for use with the Ku86:Rad54B clones



**Figure 4.** A Rad54B deficiency delays the cell death caused by the loss of function of Ku86. Five thousands cells were plated on day 0. Starting day 1, the cells were incubated in 5 nM of 4-OHT-containing media. All the cell lines utilized in this analysis express CreER. Ku86-19 and #91-19, #91-37 are sub-clones of  $Ku86^{flox/-;CreER}$  and  $Ku86^{flox/-;CreER};Rad54B^{-/-}$ #91, respectively.

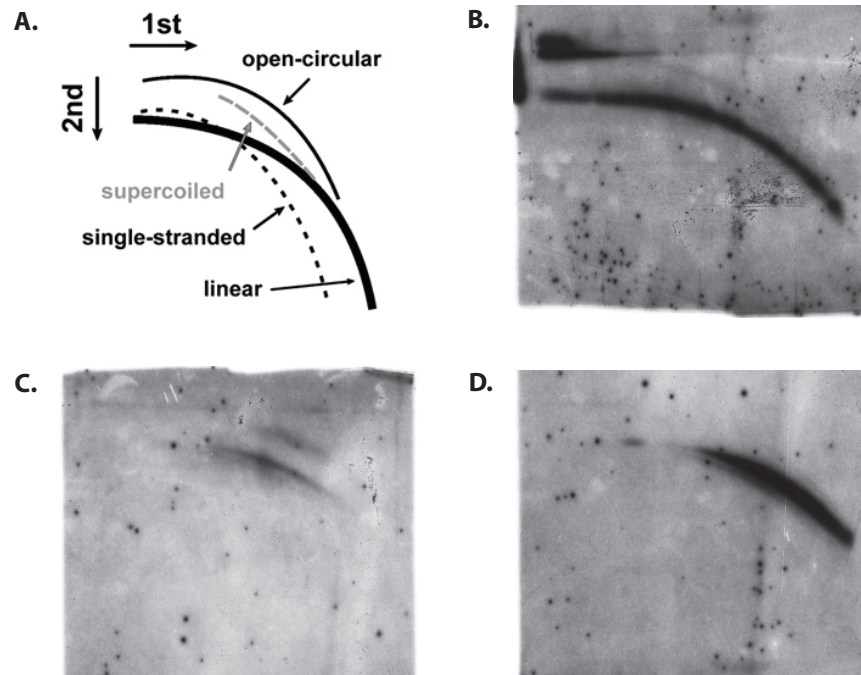
Figure 4. A Rad54B deficiency attenuates the lethality of Ku86 loss of function



**Figure 5.** A significant reduction of t-circle formation in  $Ku86^{-/-};Rad54B^{-/-}$  cells compared to  $Ku86^{-/-}$  cells. Neutral/neutral 2D-gel electrophoresis was used to detect t-circle formation. (A) A diagram showing the different arced migrations and the corresponding DNA structures. (B) Uninfected ALT WI-38 VA13 cells (C) AdCre-infected  $Ku86^{fllox/-}$  cells (D) AdCre-infected  $Ku86^{fllox/-};Rad54B^{-/-}$  #91 cells.

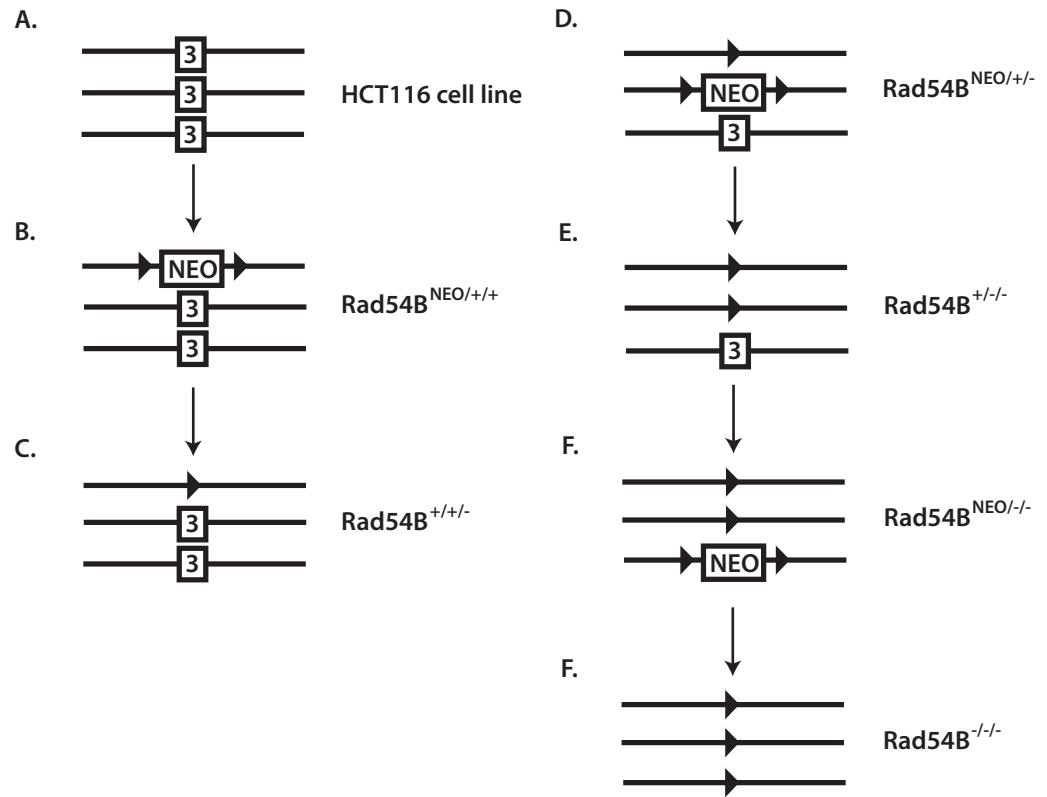


Figure 5. Neutral/neutral 2D-gel electrophoresis confirms reduced t-circle formation



**Supplementary Figure 1.** The targeting strategy and primer sets used for three rounds of Rad54B gene targeting. Primer sets #1 and #2, described in Figure 1A, were used for each round of screening. (A) The parental HCT116 cell line has three copies of the Rad54B gene and we targeted exon 3. (B) After the first round of targeting, correctly targeted clones produce a new PCR product using primer set #2. (C) Subsequent AdCre infection excises the NEO selection cassette from the Rad54B<sup>NEO/+</sup> cell line. The excision of the NEO cassette was first verified by G418-sensitivity and the G418-sensitive clones were subjected to PCR with primer set #1, and correct clones were confirmed by the absence of a PCR product from primer set #2 and the presence of new smaller product using primer set #1. Since the primer set #1 anneals outside of exon 3, the loxP-only allele produces a PCR product that is smaller than that produced from the intact exon 3. (D) The same targeting vector was used for the second round of gene targeting. All correctly targeted clones produced a PCR product using primer set #2. If the targeting vector integrated into the already inactivated allele, then the small PCR product using primer set #1 would disappear. Thus, clones that generated both small and large PCR products from primer set #1 were Rad54B<sup>NEO+/-</sup>. (E) Rad54B<sup>NEO+/-</sup> cells were infected with AdCre again and G418-sensitive clones (i.e., Rad54B<sup>+/-</sup>) were isolated and then subjected to a last round of targeting. (F) Rad54B<sup>NEO/-</sup> clones were confirmed by the presence of a PCR product using primer set #2 and the absence of the small, wild-type allele PCR product using primer set #1. (G) A final treatment with AdCre resulted in the desired genotype, Rad54B<sup>-/-</sup>.

**Supplementary Figure 1. Targeting procedure for triploid Rad54B gene**



**Table 1. Summary of Rad54B gene targeting rate in LIGIV-null and Ku86-conditional null cell lines**

**A. LIGIV-null HCT116**

Targeted Cell Type	G418 Positive	Correct Targeting	Targeting Rate (%)	Retargeted Clone
Rad54B <sup>+/+/+</sup>	34	3	8.8	–
Rad54B <sup>+/+/-</sup>	91	12	13.2	8
Rad54B <sup>+/-/-</sup>	232	11	4.7	10 *

\* one LIGIV<sup>-/-</sup>:Rad54B<sup>+/-/-</sup> clone died within a month.

**B. Ku86 conditional-null HCT116**

Targeted Cell Type	G418 Positive	Correct Targeting	Targeting Rate (%)	Retargeted Clone
Rad54B <sup>+/+/+</sup>	70	3	4.3	–
Rad54B <sup>+/+/-</sup>	229	12	5.2	5
Rad54B <sup>+/-/-</sup>	121	7	5.8	2

# **CHAPTER V**

## **Conclusions and future directions**

## Conclusions

We successfully inactivated the *LIGIV* gene in the human somatic cell line, HCT116. Prior to my work, it had been suggested that Ku86 has roles in both C-NHEJ and telomere maintenance, whereas *LIGIV* functions exclusively in C-NHEJ. Because Ku86-null cells died (117, 124) and *LIGIV*-null cells were viable (207), it had been proposed that telomere maintenance was the essential function whereas C-NHEJ activity was dispensable. While this argument is superficially compelling, the fact that different cell lines had been used to generate the relevant data was scientifically off-putting. Thus, all of the Ku86-null work had been carried out in the HCT116 colorectal carcinoma cell line while the *LIGIV*-null work had been carried out in the EBV-immortalized B-cell line, NALM-6. These differences were compounded by the fact that the NALM-6 cell line was preferentially known to use HR (357) and by the glaring world-wide absence of *LIGIV*-null patients. To address these issues, we decided to make a *LIGIV* knockout mutation in the same genetic background as the Ku86 mutations had been made; namely, the HCT116 cell line. So, we utilized rAAV-mediated gene targeting methodologies to functionally inactivate the *LIGIV* gene in HCT116 cells to verify or discredit our hypothesis. Disequilibrium during the second round of gene targeting and extreme sensitivity of *LIGIV*-null cells to DNA damaging agents vividly demonstrated the disadvantage of being *LIGIV* deficient. Nonetheless, we successfully generated a viable *LIGIV*-null cell line. Although our data do not address why *LIGIV*-null patients have not been identified, they do show unequivocally that human somatic cells do not require *LIGIV*.

In a related project, we demonstrated that the functional inactivation of *LIGIII* results in cell death. While theoretically provocative, the practical limitations of performing biochemical assays with dead cells were daunting. To obviate this biological obstacle, we next successfully generated a viable nuclear *LIGIII*-null HCT116 cell line using mitochondrial-specific complementation. Although a demonstration that *LIGIII* is needed for mitochondrial maintenance was important, we were inherently more interested in *LIGIII*'s nuclear roles. Unfortunately, the absence of *LIGIII* from the nucleus was phenotypically unnoticeable (with one exception to be described below). Perhaps most disappointing was the absence of an effect on rAAV-mediated gene targeting, where

work from several groups (327, 378), including ours (276), had strongly implicated LIGIII. Although much of the data generated using the conditionally-null LIGIII cell line was negative, the construction of the cell line was important in and of its own right since it is the first human cell line ever described with the loss-of-function of a key candidate A-NHEJ pathway gene.

Finally, we constructed a  $Ku86^{flox/-};Rad54B^{-/-}$  cell line.  $Ku86$ -null cells were known to die due to telomere defects including massive telomere loss, sister chromatid fusions and t-circle formation (117). Work from our laboratory and other laboratories had circumstantially implicated aberrant HR activity as being responsible for the t-circle formation. To scientifically address this issue, we attempted to functionally inactivate the HR pathway in the conditionally-null  $Ku86$  cell line. To this end, we inactivated the  $Rad54B$  gene, because  $Rad54B^{-/-}$  cells were known to be HR defective based upon their inability to undergo further gene targeting (83). In a rare instance of experimental results corresponding to theoretical predictions, we observed that  $Ku86^{-/-};Rad54B^{-/-}$  cells show significantly reduced t-circle formation compared to  $Ku86^{-/-}$  cells. In addition,  $Ku86^{-/-};Rad54B^{-/-}$  cells survived longer. Both of these observations suggested that a  $Rad54B$  deficiency can partially rescue the pathologies associated with  $Ku86$ -null cells. More specifically, our data are the first good genetic data implicating HR in the genesis of t-circles.

## Future Directions

**rAAV-mediated gene targeting is a magic bullet for investigators working with human somatic cell lines.**

In the course of my thesis work, I utilized rAAV-mediated gene targeting technology to functionally inactivate three genes ( $LIGIV$ ,  $LIGIII$  and  $RAD54B$ ) that are required for each of the three major DNA DSB repair pathways (C-NHEJ, A-NHEJ and HR, respectively). The utility of having cell lines specifically defective in a particular DNA DSB repair pathway is evident from 1) the data presented in this thesis, 2) the fact that these cell lines have already found their way into published works (117, 276, 363), and from their application in commercial endeavors (unpublished observations). Moreover, we have also already demonstrated our ability to establish second-generation cell lines,

which contain genetic alterations at two different loci (e.g., Ku86<sup>fllox/-</sup>:Rad54B<sup>-/-</sup> cells). The rAAV methodology has been designed to be “scarless”, *i.e.*, such that no drug selection markers are left in the chromosome after the targeting event. Because of this feature, the methodology can, in principle, be limitlessly applied to a single cell line — making as many genomic alterations as the investigator wishes. This ability, once the sole purview of bacterial or yeast geneticists, has greatly enhanced the attractiveness and utility of human somatic cells as a valid model system. In addition, it should be noted that we have confined ourselves to mostly making gross alterations (“knockouts”) of the human genome. The existing rAAV methodology is, however, much more refined and established vectors and protocols already exist for the use of rAAV to make: 1) conditional alleles, 2) knock-in mutations and inserting epitope tags at the 3) N- and 4) C-termini of chromosomal genes (117, 330). This powerful methodology makes the investigator the limiting factor for answering most biological questions. There is no reason (and indeed many of these experiments are already ongoing in our laboratory) that the cell lines that I have generated during my thesis work, can not be further modified to address a myriad of questions in DNA repair, DNA replication, DNA recombination and telomere maintenance.

### **Introducing an inducible-CreER systems an improvement to study conditionally-null cell lines**

Our laboratory has already generated two conditionally-null cell lines (Ku86<sup>fllox/-</sup> and LIG3<sup>fllox/-</sup>) to study essential genes {(117); and chapter 3} and is in the process of making many more. All of these constructions utilize the technology of Cre:LoxP site-specific recombination (414). At some step — and often at many steps — the expression of the Cre recombinase is required to mediate the appropriate recombination event. In the past, our laboratory has used plasmids, adenoviruses, lentiviruses and hit-and-run/suicidal gene expression constructs to express Cre. Unfortunately, all of these methods have associated problems including high toxicity, or too low or too high of an efficiency. These problems have varied in their severity from mildly irritating to so severe that the desired experiment could not be accomplished. Here, for the first time, I made use of the 4-OHT inducible Cre recombinase system following its introduction by PiggyBac-mediated



transposition. This system shows almost no toxicity and it is easy to control (Chapter 5), making it a major improvement to the existing technologies. We predict that introducing the inducible CreER system into the  $LIGIII^{flox/-}$  and future conditionally-null cell lines will solve a lot of attendant technical problems and will become a generally valuable tool to the field to study essential genes.

### **Future uses of the conditionally-null LIGIII cell line: 1) generation of an A-NHEJ chromosomal translocation reporter system**

Perhaps the single biggest surprise (and disappointment) of my thesis work was the lack of phenotypes associated with the loss-of-function of nuclear LIGIII (Chapter 4). Significant amounts of circumstantial evidence had suggested that LIGIII was required for rAAV-mediated integrations and for micro-homology mediated end-joining. Alas, no defects in either of those processes were observed. Nonetheless, we remain convinced that the construction of the LIGIII conditionally-null cell line will prove itself useful. First, based on the sequencing of chromosomal translocations in human cancer cell lines, it had been inferred that LIGIII/A-NHEJ might be involved in their genesis, since the junctions often contained microhomology. That hypothesis gained significant support from analyses of chromosomal translocations in conditionally-null LIGIII MEFs (272). We think we can extend these observations (which have a direct impact on the initiation of tumorigenesis) in our cell line. Thus, we propose to introduce *I-SceI* restriction sites at two specific chromosomal locations using rAAV-mediated gene targeting technology. *I-SceI* is a restriction enzyme whose recognition sequence is not normally found in the human genome (415). Once this cell line is established, *I-SceI* can be transiently expressed and chromosomal translocations can be detected by FISH using whole chromosome painting techniques (380, 416, 417). A more complicated variation of this experiment would be to imbed two halves of a selectable marker (*e.g.*, GFP, LacZ, NEO, *etc.*) adjacent to the *I-SceI* sites (416, 417). In this scenario, the selectable marker is reconstituted only when *I-SceI*-induced chromosomal translocations occur (272, 380). Other methodologies could also be used. For example, a pair of zinc-finger nucleases could be used to induce DSBs at two specific loci. If chromosomal translocations happened, they could be identified and isolated using PCR approaches where one of the

primers hybridizes adjacent to one of the zinc-finger nuclease binding sites and another primer hybridizes next to the other site. Similar approaches have been described using mouse cell lines (272). Using one of these approaches we think we could use our cell line to study the impact of the presence and absence of LIGIII expression on chromosomal translocations in human cells.

**Future uses of the conditionally-null LIGIII cell line: 2) is LIGIII required to bypass the crisis induced by gradual telomere loss?**

Although it was not the focus of my thesis work, one of the tangential issues I was interested in was to study the mechanism of telomere fusions in Ku86-deficient cells. We have described extensively our work looking at t-circle formation associated with the loss of Ku86 expression (Chapter 5). One of the other features of the Ku86 loss-of-function mutation is that it induces sister telomere fusions (117). This feature was in fact the motivation for generating the Ku86<sup>fllox/-</sup>:LIGIV<sup>-/-</sup> cell line. These cells, however, showed no impairment in sister chromatid fusions in comparison to the Ku86<sup>fllox/-</sup> cell line (data not shown) demonstrating a lack of involvement of LIGIV in this process. This observation, in turn, was one of the most relevant rationales for generating the LIGIII<sup>fllox/-</sup> cell line (Chapter 4). Indeed, the construction of a Ku86<sup>fllox/-</sup>:LIGIII<sup>fllox/-</sup> cell line, with which we intend to study sister:sister chromatid fusion is almost completed (data not shown). In the interim, however, we decided to see if we could extend the utility of our mutant cell lines. Thus, the loss of telomeres mediated by Ku86 loss-of-function, while intellectually interesting, is nonetheless not very relevant to normal cells. In normal cells, the relevant telomere loss is the slow, gradual reduction of the telomere due to cellular proliferation, which, if left unchecked can lead to the progression of malignancy (418). Therefore we initiated collaboration with the laboratory of Duncan Baird to analyze the impact of our loss-of-function mutant cell lines on gradual telomere loss. To this end, Dr. Baird's laboratory established stable derivatives of our wild-type HCT116, LIGIV<sup>-/-</sup> and LIGIII<sup>-/-</sup> cells using a DN-hTERT construct that reduces telomerase activity. Dr. Baird's laboratory had shown that the expression of the DN-hTERT construct in other human cell lines over the course of several months resulted in a gradual telomere shortening that ultimately led to telomere fusions and a "crisis"-like state (419). In our cell lines,

telomere shortening occurred and telomere fusions could be detected as the cells entered a crisis-like state (data not shown). These results implied that neither LIGIV nor LIGIII functions in the pathway of telomere fusions resulting from gradual telomere shortening. Subsequently, however, we made a startling preliminary observation. After entering crisis, eleven of eleven independent wild-type subclones and four of four independent LIGIV<sup>-/-</sup> subclones recovered and multiple survivor clones grew out. In striking contrast, none of eleven independent LIG3<sup>-/-</sup> sub-clones yielded a single survivor. These preliminary data were supported by knockdown experiments in other human cell lines where it was observed that a reduction in LIGIV had no effect on telomere fusions or survival whereas LIGIII knockdown cells yielded chromosome fusions, but not survivors (data not shown). These studies, albeit preliminary and having been performed only once, indicate the presence of a LIGIII-dependent event that is required for human cells to survive the crisis induced by gradual telomere shortening.

To further characterize this phenomenon, we will sequence the telomeric fusion junctions produced from each cell line in the hope that the molecular signatures at the repaired junctions will give us some insight into whether certain junctions accumulate or are absent in the LIGIII-null cells. Additional genetic experiments are also planned. A complementation experiment where a nuclear LIGIII expression construct has been introduced into the LIGIII conditionally-null cell line is already underway. Additional preliminary data from Dr. Baird's laboratory has implicated p53 in this phenomenon. Thus, we plan to knockout the p53 gene in the LIGIII<sup>-/-mL3</sup> cell line to determine whether the failure of LIGIII<sup>-/-</sup> cells to survive crisis is mediated by p53. Since virtually all human cancer cells are postulated to arise only after surviving crisis, we feel that these experiments have the possibility to have important clinical ramifications.

### **Future uses of the conditionally-null LIGIII cell line: 3) structure:function analyses of LIGIII**

An additional mundane, albeit important, use for the conditionally-null LIGIII cell line would be to use it as a tool in classic structure:function (*aka*, "wreck and check") experiments. Thus, LIGIII contains several readily-identifiable domains: a zinc-finger domain, a BRCT domain, etc. but the functional significance of these domains remains

poorly understood since there has been no experimental system available to dissect their function. Our conditionally-null LIGIII cell line fills that scientific void. Thus, (assuming we ultimately have a phenotype or biological endpoint that we can quantitatively measure) wild-type nuclear LIGIII expression constructs containing mutations in these domains can be expressed in the cells and their impact assessed. While not hypothesis-driven *per se*, this practical approach to assessing a gene's function(s) has historically proven to be very useful. In the same vein, we can use the LIGIV- and RAD54B-null cell lines for an identical purpose.

## **Interactions between the HR, C-NHEJ and A-NHEJ pathways**

### **1) Pathway choice mechanism**

The initial end resection process of A-NHEJ is similar to that of HR. Unlike HR, however, there is no evidence that A-NHEJ requires EXO1 and DNA2, the nucleases required for the several kilobases of end resection that occur in HR (54, 90). This suggests that the degree of resection may differentiate between whether a DSB is repaired via HR or A-NHEJ. A-NHEJ is most active in G2 phase and its activity is compromised when cells enter the plateau phase of growth, which was demonstrated in mouse MEF and hamster CHO cell lines (285-287). Thus, the cell cycle dependent regulation of nucleases might be one way in which pathway choice is determined (45, 288, 290). To understand if pathway choice is mediated by end resection, it will be necessary to generate cell lines deficient in the nucleases involved in early steps of DSB repair, including Mre11, CtIP, Exo1 and Dna2 (22, 34, 37, 54, 87-90). An analysis of the relative activities of HR and A-NHEJ in these cell lines should ensure a better understanding of repair pathway choice mechanisms.

We could enhance these studies with the use of modified repair substrates to determine how the degree of end resection influences repair pathway choice. In a C-NHEJ-deficient condition, repair substrates with microhomology positioned at different intervals from a predetermined DSB could be used to test how much resection is required for A-NHEJ.

### **2) Inactivation of other A-NHEJ components**

The *LIGIII*<sup>-/-</sup> HCT116 cell line we have described herein (Chapter 3) is the only available A-NHEJ-deficient human cell line. Inactivation of other A-NHEJ components will be necessary to understand the A-NHEJ pathway in human cells. Our laboratory has already produced a complete series — encompassing seven different genes — of C-NHEJ-deficient cell lines using rAAV-mediated gene targeting as a proof of principle that this is a viable way to approach this problem. Thus, one of the future goals of the laboratory is to make knockout cell lines corresponding to the other genes that have been implicated in A-NHEJ: MRN, PARP-1, XRCC1, CtIP and histone H1 (27, 44, 248, 250, 252, 420). Indeed, we have already made a *PARP-1*<sup>-/-</sup> cell line (M. Mueller and E. A. Hendrickson, unpublished) and other constructions are ongoing.

### **3) What is the minimum DSB repair activity required for cell survival?**

We have already observed that combined C-NHEJ:A-NHEJ or HR:C-NHEJ deficiencies, respectively, cause cell death. In particular, it is interesting that the deficiency of C-NHEJ and A-NHEJ results in synthetic lethality, because it suggests that A-NHEJ and C-NHEJ are not completely redundant, but that A-NHEJ probably has its own (albeit unknown), distinct functions. This non-redundancy was hinted at previously, when it was shown that *LIGI* and *LIGIII*, but not *LIGIV*, were required for MMEJ (421). Inactivation of other A-NHEJ components should help us assess the unique function of A-NHEJ. Finally, it will be also interesting to characterize the viability of HR and A-NHEJ deficient cell lines, because it should help us determine whether HR and A-NHEJ are redundant for certain tasks, as well as addressing what the minimum DNA DSB repair activity is that is required for cell survival.

## Bibliography

1. Norbury CJ, Zhivotovsky B (2004) DNA damage-induced apoptosis. *Oncogene* 23:2797–2808.
2. Lieber MR (2010) The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annu Rev Biochem*.
3. Gellert M (1994) DNA double-strand breaks and hairpins in V(D)J recombination. *Semin. Immunol.* 6:125–130.
4. Kenter AL (2005) Class switch recombination: an emerging mechanism. *Curr. Top. Microbiol. Immunol.* 290:171–199.
5. Hendrickson EA et al. (1991) A link between double-strand break-related repair and V(D)J recombination: the scid mutation. *Proc Natl Acad Sci USA* 88:4061–4065.
6. Yanowitz J (2010) Meiosis: making a break for it. *Curr. Opin. Cell Biol.* 22:744–751.
7. Petermann E, Helleday T (2010) Pathways of mammalian replication fork restart. *Nat Rev Mol Cell Biol* 11:683–687.
8. Sancar A, Lindsey-Boltz LA, Unsal-Kaçmaz K, Linn S (2004) Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu Rev Biochem* 73:39–85.
9. Kanaar R, Hoeijmakers JH, van Gent DC (1998) Molecular mechanisms of DNA double strand break repair. *Trends Cell Biol.* 8:483–489.
10. Wyman C, Ristic D, Kanaar R (2004) Homologous recombination-mediated double-strand break repair. *DNA Repair (Amst)* 3:827–833.
11. Takata M et al. (1998) Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J* 17:5497–5508.
12. Essers J, Essers J (2000) Homologous and non-homologous recombination differentially affect DNA damage repair in mice. *EMBO J* 19:1703–1710.
13. Li X, Heyer W-D (2008) Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res* 18:99–113.
14. Huertas P (2010) DNA resection in eukaryotes: deciding how to fix the break. *Nat Struct Mol Biol* 17:11–16.
15. Daley JM, Palmbo PL, Wu D, Wilson TE (2005) Nonhomologous end joining in yeast. *Annu Rev Genet* 39:431–451.
16. Shrivastav M, De Haro LP, Nickoloff JA (2008) Regulation of DNA double-strand break repair pathway choice. *Cell Res* 18:134–147.
17. Honma M (2005) Generation of loss of heterozygosity and its dependency on p53 status in human lymphoblastoid cells. *Environ. Mol. Mutagen.* 45:162–176.
18. Hefferin ML, Tomkinson AE (2005) Mechanism of DNA double-strand break repair by non-homologous end joining. *DNA Repair (Amst)* 4:639–

648.

19. Lieber MR (2008) The mechanism of human nonhomologous DNA end joining. *J Biol Chem* 283:1–5.
20. Ma J-L, Kim EM, Haber JE, Lee SE (2003) Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. *Mol Cell Biol* 23:8820–8828.
21. Rothkamm K, Krüger I, Thompson LH, Löbrich M (2003) Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol Cell Biol* 23:5706–5715.
22. You Z, Bailis JM (2010) DNA damage and decisions: CtIP coordinates DNA repair and cell cycle checkpoints. *Trends Cell Biol.* 20:402–409.
23. Verkaik NS et al. (2002) Different types of V(D)J recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells. *Eur J Immunol* 32:701–709.
24. Kabotyanski EB, Gomelsky L, Han JO, Stamato TD, Roth DB (1998) Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Res* 26:5333–5342.
25. Chang C, Biedermann KA, Mezzina M, Brown JM (1993) Characterization of the DNA double strand break repair defect in scid mice. *Cancer Res* 53:1244–1248.
26. Tomlinson CG et al. (2011) Neutralizing mutations of carboxylates that bind metal 2 in t5 flap endonuclease result in an enzyme that still requires two metal ions. *J Biol Chem* 286:30878–30887.
27. Wang M et al. (2006) PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res* 34:6170–6182.
28. Khanna KK, Jackson SP (2001) DNA double-strand breaks: signaling, repair and the cancer connection. *Nat Genet* 27:247–254.
29. Durocher D, Jackson SP (2001) DNA-PK, ATM and ATR as sensors of DNA damage: variations on a theme? *Curr. Opin. Cell Biol.* 13:225–231.
30. Lovejoy CA, Cortez D (2009) Common mechanisms of PIKK regulation. *DNA Repair (Amst)* 8:1004–1008.
31. Hiom K (2005) DNA repair: how to PIKK a partner. *Curr Biol* 15:R473–5.
32. Yamamoto K-I, Kobayashi M, Shimizu H (2006) ATM, a paradigm for a stress-responsive signal transducer in higher vertebrate cells. *Subcell. Biochem.* 40:327–339.
33. Matsuoka S et al. (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* 316:1160–1166.
34. Jazayeri A et al. (2006) ATM- and cell cycle-dependent regulation of ATR in response to DNA double-strand breaks. *Nat Cell Biol* 8:37–45.
35. Falck J, Coates J, Jackson SP (2005) Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature* 434:605–611.
36. Helmink BA et al. (2009) MRN complex function in the repair of chromosomal Rag-mediated DNA double-strand breaks. *J Exp Med*

- 206:669–679.
37. Dupré A, Boyer-Chatenet L, Gautier J (2006) Two-step activation of ATM by DNA and the Mre11-Rad50-Nbs1 complex. *Nature Publishing Group* 13:451–457.
  38. Unsal-Kaçmaz K, Sancar A (2004) Quaternary structure of ATR and effects of ATRIP and replication protein A on its DNA binding and kinase activities. *Mol Cell Biol* 24:1292–1300.
  39. So S, Davis AJ, Chen DJ (2009) Autophosphorylation at serine 1981 stabilizes ATM at DNA damage sites. *J Cell Biol* 187:977–990.
  40. Zou L, Elledge SJ (2003) Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* 300:1542–1548.
  41. Cortez D, Guntuku S, Qin J, Elledge SJ (2001) ATR and ATRIP: partners in checkpoint signaling. *Science* 294:1713–1716.
  42. Stiff T et al. (2006) ATR-dependent phosphorylation and activation of ATM in response to UV treatment or replication fork stalling. *EMBO J* 25:5775–5782.
  43. Lamarche BJ, Orazio NI, Weitzman MD (2010) The MRN complex in double-strand break repair and telomere maintenance. *FEBS Lett* 584:3682–3695.
  44. Zhuang J, Jiang G, Willers H, Xia F (2009) Exonuclease function of human Mre11 promotes deletional nonhomologous end joining. *J Biol Chem* 284:30565–30573.
  45. Zhu XD, Küster B, Mann M, Petrini JH, de Lange T (2000) Cell-cycle-regulated association of RAD50/MRE11/NBS1 with TRF2 and human telomeres. *Nat Genet* 25:347–352.
  46. Chai W, Sfeir AJ, Hoshiyama H, Shay JW, Wright WE (2006) The involvement of the Mre11/Rad50/Nbs1 complex in the generation of G-overhangs at human telomeres. *EMBO Rep* 7:225–230.
  47. O'Sullivan RJ, Karlseder J (2010) Telomeres: protecting chromosomes against genome instability. *Nat Rev Mol Cell Biol* 11:171–181.
  48. Lombard DB, Guarente L (2000) Nijmegen breakage syndrome disease protein and MRE11 at PML nuclear bodies and meiotic telomeres. *Cancer Res* 60:2331–2334.
  49. Williams GJ, Lees-Miller SP, Tainer JA (2010) Mre11-Rad50-Nbs1 conformations and the control of sensing, signaling, and effector responses at DNA double-strand breaks. *DNA Repair (Amst)* 9:1299–1306.
  50. Hopfner KP et al. (2001) Structural biochemistry and interaction architecture of the DNA double-strand break repair Mre11 nuclease and Rad50-ATPase. *Cell* 105:473–485.
  51. Williams RS et al. (2008) Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair. *Cell* 135:97–109.
  52. Buis J et al. (2008) Mre11 nuclease activity has essential roles in DNA repair and genomic stability distinct from ATM activation. *Cell* 135:85–96.



53. Taylor EM et al. (2010) The Mre11/Rad50/Nbs1 complex functions in resection-based DNA end joining in *Xenopus laevis*. *Nucleic Acids Res* 38:441–454.
54. Farah JA, Cromie GA, Smith GR (2009) Ctp1 and Exonuclease 1, alternative nucleases regulated by the MRN complex, are required for efficient meiotic recombination. *Proceedings of the National Academy of Sciences* 106:9356–9361.
55. Hopfner K-P et al. (2002) The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. *Nature* 418:562–566.
56. Difilippantonio S et al. (2005) Role of Nbs1 in the activation of the Atm kinase revealed in humanized mouse models. *Nat Cell Biol* 7:675–685.
57. Carney JP et al. (1998) The hMre11/hRad50 protein complex and Nijmegen breakage syndrome: linkage of double-strand break repair to the cellular DNA damage response. *Cell* 93:477–486.
58. Antoccia A, Kobayashi J, Tauchi H, Matsuura S, Komatsu K (2006) Nijmegen breakage syndrome and functions of the responsible protein, NBS1. *Genome Dyn* 1:191–205.
59. Waltes R et al. (2009) Human RAD50 deficiency in a Nijmegen breakage syndrome-like disorder. *Am. J. Hum. Genet.* 84:605–616.
60. Lee J-H, Lee J-H, Paull TT, Paull TT (2007) Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. *Oncogene* 26:7741–7748.
61. Martinez-Perez E, Colaiácovo MP (2009) Distribution of meiotic recombination events: talking to your neighbors. *Curr. Opin. Genet. Dev.* 19:105–112.
62. Tsang E, Carr AM (2008) Replication fork arrest, recombination and the maintenance of ribosomal DNA stability. *DNA Repair (Amst)* 7:1613–1623.
63. Lin Z, Kong H, Nei M, Ma H (2006) Origins and evolution of the recA/RAD51 gene family: evidence for ancient gene duplication and endosymbiotic gene transfer. *Proc Natl Acad Sci USA* 103:10328–10333.
64. Suwaki N, Klare K, Tarsounas M (2011) RAD51 paralogs: Roles in DNA damage signalling, recombinational repair and tumorigenesis. *Semin Cell Dev Biol.*
65. Sugiyama T, Kantake N, Wu Y, Kowalczykowski SC (2006) Rad52-mediated DNA annealing after Rad51-mediated DNA strand exchange promotes second ssDNA capture. *EMBO J* 25:5539–5548.
66. Mortensen UH, Bendixen C, Sunjevaric I, Rothstein R (1996) DNA strand annealing is promoted by the yeast Rad52 protein. *Proc Natl Acad Sci USA* 93:10729–10734.
67. Mortensen UH, Lisby M, Rothstein R (2009) Rad52. *Curr Biol* 19:R676–7.
68. Rijkers T et al. (1998) Targeted inactivation of mouse RAD52 reduces homologous recombination but not resistance to ionizing radiation. *Mol Cell Biol* 18:6423–6429.

69. Singleton MR, Wentzell LM, Liu Y, West SC, Wigley DB (2002) Structure of the single-strand annealing domain of human RAD52 protein. *Proc Natl Acad Sci USA* 99:13492–13497.
70. Wooster R, Weber BL (2003) Breast and ovarian cancer. *N. Engl. J. Med.* 348:2339–2347.
71. Tutt A, Ashworth A (2002) The relationship between the roles of BRCA genes in DNA repair and cancer predisposition. *Trends Mol Med* 8:571–576.
72. Pellegrini L et al. (2002) Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* 420:287–293.
73. Yang H, Li Q, Fan J, Holloman WK, Pavletich NP (2005) The BRCA2 homologue Brh2 nucleates RAD51 filament formation at a dsDNA-ssDNA junction. *Nature* 433:653–657.
74. Wesoly J et al. (2006) Differential contributions of mammalian Rad54 paralogs to recombination, DNA damage repair, and meiosis. *Mol Cell Biol* 26:976–989.
75. Kiiianitsa K, Solinger JA, Heyer W-D (2006) Terminal association of Rad54 protein with the Rad51-dsDNA filament. *Proc Natl Acad Sci USA* 103:9767–9772.
76. Tan TLR, Kanaar R, Wyman C (2003) Rad54, a Jack of all trades in homologous recombination. *DNA Repair (Amst)* 2:787–794.
77. Mazin AV, Mazina OM, Bugreev DV, Rossi MJ (2010) Rad54, the motor of homologous recombination. *DNA Repair (Amst)* 9:286–302.
78. Li X et al. (2007) Rad51 and Rad54 ATPase activities are both required to modulate Rad51-dsDNA filament dynamics. *Nucleic Acids Res* 35:4124–4140.
79. Rossi MJ, Mazin AV (2008) Rad51 protein stimulates the branch migration activity of Rad54 protein. *J Biol Chem* 283:24698–24706.
80. Sung P, Krejci L, Van Komen S, Sehorn MG (2003) Rad51 recombinase and recombination mediators. *J Biol Chem* 278:42729–42732.
81. Tanaka K, Hiramoto T, Fukuda T, Miyagawa K (2000) A novel human rad54 homologue, Rad54B, associates with Rad51. *J Biol Chem* 275:26316–26321.
82. Tanaka K, Kagawa W, Kinebuchi T, Kurumizaka H, Miyagawa K (2002) Human Rad54B is a double-stranded DNA-dependent ATPase and has biochemical properties different from its structural homolog in yeast, Tid1/Rdh54. *Nucleic Acids Res* 30:1346–1353.
83. Miyagawa K et al. (2002) A role for RAD54B in homologous recombination in human cells. *EMBO J* 21:175–180.
84. Uziel T et al. (2003) Requirement of the MRN complex for ATM activation by DNA damage. *EMBO J* 22:5612–5621.
85. Lee J-H, Paull TT (2005) ATM activation by DNA double-strand breaks through the Mre11-Rad50-Nbs1 complex. *Science* 308:551–554.
86. Lee J-H, Paull TT (2004) Direct activation of the ATM protein kinase by the

- Mre11/Rad50/Nbs1 complex. *Science* 304:93–96.
87. Yuan J, Chen J (2009) N terminus of CtIP is critical for homologous recombination-mediated double-strand break repair. *J Biol Chem* 284:31746–31752.
  88. You Z et al. (2009) CtIP links DNA double-strand break sensing to resection. *Mol Cell* 36:954–969.
  89. Sartori AA et al. (2007) Human CtIP promotes DNA end resection. *Nature* 450:509–514.
  90. Nimonkar AV et al. (2011) BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes Dev* 25:350–362.
  91. Sugiyama T, Zaitseva EM, Kowalczykowski SC (1997) A single-stranded DNA-binding protein is needed for efficient presynaptic complex formation by the *Saccharomyces cerevisiae* Rad51 protein. *J Biol Chem* 272:7940–7945.
  92. New JH, Kowalczykowski SC (2002) Rad52 protein has a second stimulatory role in DNA strand exchange that complements replication protein-A function. *J Biol Chem* 277:26171–26176.
  93. Sugiyama T, New JH, Kowalczykowski SC (1998) DNA annealing by RAD52 protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. *Proc Natl Acad Sci USA* 95:6049–6054.
  94. Sung P, Robberson DL (1995) DNA strand exchange mediated by a RAD51-ssDNA nucleoprotein filament with polarity opposite to that of RecA. *Cell* 82:453–461.
  95. Shinagawa H, Iwasaki H (1995) Molecular mechanisms of Holliday junction processing in *Escherichia coli*. *Adv. Biophys.* 31:49–65.
  96. Rass U et al. (2010) Mechanism of Holliday junction resolution by the human GEN1 protein. *Genes Dev* 24:1559–1569.
  97. Lorenz A, West SC, Whitby MC (2010) The human Holliday junction resolvase GEN1 rescues the meiotic phenotype of a *Schizosaccharomyces pombe* mus81 mutant. *Nucleic Acids Res* 38:1866–1873.
  98. West SC (2009) The search for a human Holliday junction resolvase. *Biochem. Soc. Trans.* 37:519–526.
  99. Wechsler T, Newman S, West SC (2011) Aberrant chromosome morphology in human cells defective for Holliday junction resolution. *Nature* 471:642–646.
  100. Lieber MR, Ma Y, Pannicke U, Schwarz K (2004) The mechanism of vertebrate nonhomologous DNA end joining and its role in V(D)J recombination. *DNA Repair (Amst)* 3:817–826.
  101. Fugmann SD, Lee AI, Shockett PE, Villey IJ, Schatz DG (2000) The RAG proteins and V(D)J recombination: complexes, ends, and transposition. *Annu Rev Immunol* 18:495–527.

102. McBride KM et al. (2008) Regulation of class switch recombination and somatic mutation by AID phosphorylation. *J Exp Med* 205:2585–2594.
103. Di Noia JM, Neuberger MS (2007) Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem* 76:1–22.
104. Kotnis A, Du L, Liu C, Popov SW, Pan-Hammarström Q (2009) Non-homologous end joining in class switch recombination: the beginning of the end. *Philos Trans R Soc Lond, B, Biol Sci* 364:653–665.
105. Lange T de (2010) Telomere biology and DNA repair: Enemies with benefits. *FEBS Lett* 584:3673–3674.
106. Diotti R, Loayza D (2011) Shelterin complex and associated factors at human telomeres. *Nucleus* 2:119–135.
107. de Lange T (2005) Shelterin: the protein complex that shapes and safeguards human telomeres. *Genes Dev* 19:2100–2110.
108. Palm W, de Lange T (2008) How Shelterin Protects Mammalian Telomeres. *Annu Rev Genet* 42:301–334.
109. Wu P, van Overbeek M, Rooney S, de Lange T (2010) Apollo contributes to G overhang maintenance and protects leading-end telomeres. *Mol Cell* 39:606–617.
110. Williams ES et al. (2009) Telomere dysfunction and DNA-PKcs deficiency: characterization and consequence. *Cancer Res* 69:2100–2107.
111. Riha K, Watson JM, Parkey J, Shippen DE (2002) Telomere length deregulation and enhanced sensitivity to genotoxic stress in Arabidopsis mutants deficient in Ku70. *EMBO J* 21:2819–2826.
112. Hsu HL et al. (2000) Ku acts in a unique way at the mammalian telomere to prevent end joining. *Genes Dev* 14:2807–2812.
113. Hsu HL, Gilley D, Blackburn EH, Chen DJ (1999) Ku is associated with the telomere in mammals. *Proc Natl Acad Sci USA* 96:12454–12458.
114. Indiviglio SM, Bertuch AA (2009) Ku's essential role in keeping telomeres intact. *Proceedings of the National Academy of Sciences* 106:12217–12218.
115. Myung K et al. (2004) Regulation of telomere length and suppression of genomic instability in human somatic cells by Ku86. *Mol Cell Biol* 24:5050–5059.
116. Fisher TS, Zakian VA (2005) Ku: a multifunctional protein involved in telomere maintenance. *DNA Repair (Amst)* 4:1215–1226.
117. Wang Y, Ghosh G, Hendrickson EA (2009) Ku86 represses lethal telomere deletion events in human somatic cells. *Proceedings of the National Academy of Sciences* 106:12430–12435.
118. Zellinger B, Akimcheva S, Puizina J, Schirato M, Riha K (2007) Ku Suppresses Formation of Telomeric Circles and Alternative Telomere Lengthening in Arabidopsis. *Mol Cell* 27:163–169.
119. d'Adda di Fagagna F et al. (2001) Effects of DNA nonhomologous end-joining factors on telomere length and chromosomal stability in mammalian cells. *Curr Biol* 11:1192–1196.

120. de Lange T (2009) How Telomeres Solve the End-Protection Problem. *Science* 326:948–952.
121. Mimori T et al. (1981) Characterization of a high molecular weight acidic nuclear protein recognized by autoantibodies in sera from patients with polymyositis-scleroderma overlap. *J Clin Invest* 68:611–620.
122. Errami A et al. (1996) Ku86 defines the genetic defect and restores X-ray resistance and V(D)J recombination to complementation group 5 hamster cell mutants. *Mol Cell Biol* 16:1519–1526.
123. Taccioli GE et al. (1994) Ku80: product of the XRCC5 gene and its role in DNA repair and V(D)J recombination. *Science* 265:1442–1445.
124. Li G, Nelsen C, Hendrickson EA (2002) Ku86 is essential in human somatic cells. *Proc Natl Acad Sci USA* 99:832–837.
125. Douglas P, Gupta S, Morrice N, Meek K, Lees-Miller SP (2005) DNA-PK-dependent phosphorylation of Ku70/80 is not required for non-homologous end joining. *DNA Repair (Amst)* 4:1006–1018.
126. Mimori T, Hardin JA (1986) Mechanism of interaction between Ku protein and DNA. *J Biol Chem* 261:10375–10379.
127. Pang D, Yoo S, Dynan WS, Jung M, Dritschilo A (1997) Ku proteins join DNA fragments as shown by atomic force microscopy. *Cancer Res* 57:1412–1415.
128. Arosio D et al. (2002) Studies on the mode of Ku interaction with DNA. *J Biol Chem* 277:9741–9748.
129. Walker JR, Corpina RA, Goldberg J (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* 412:607–614.
130. Critchlow SE, Jackson SP (1998) DNA end-joining: from yeast to man. *Trends in Biochemical Sciences* 23:394–398.
131. Gell D, Jackson SP (1999) Mapping of protein-protein interactions within the DNA-dependent protein kinase complex. *Nucleic Acids Res* 27:3494–3502.
132. Singleton BK, Torres Arzayus MI, Rottinghaus ST, Taccioli GE, Jeggo PA (1999) The C terminus of Ku80 activates the DNA-dependent protein kinase catalytic subunit. *Mol Cell Biol* 19:3267–3277.
133. Mahajan KN, Nick McElhinny SA, Mitchell BS, Ramsden DA (2002) Association of DNA polymerase mu (pol mu) with Ku and ligase IV: role for pol mu in end-joining double-strand break repair. *Mol Cell Biol* 22:5194–5202.
134. Ma Y et al. (2004) A biochemically defined system for mammalian nonhomologous DNA end joining. *Mol Cell* 16:701–713.
135. Nick McElhinny SA, Snowden CM, McCarville J, Ramsden DA (2000) Ku recruits the XRCC4-ligase IV complex to DNA ends. *Mol Cell Biol* 20:2996–3003.
136. Mari P-O et al. (2006) Dynamic assembly of end-joining complexes requires

- interaction between Ku70/80 and XRCC4. *Proc Natl Acad Sci USA* 103:18597–18602.
137. Palmbos PL, Daley JM, Wilson TE (2005) Mutations of the Yku80 C terminus and Xrs2 FHA domain specifically block yeast nonhomologous end joining. *Mol Cell Biol* 25:10782–10790.
138. Hsu HL, Yannone SM, Chen DJ (2002) Defining interactions between DNA-PK and ligase IV/XRCC4. *DNA Repair (Amst)* 1:225–235.
139. Lopez CR et al. (2011) Ku Must Load Directly onto the Chromosome End in Order to Mediate Its Telomeric Functions. *PLoS Genet* 7:e1002233.
140. Ribes-Zamora A, Mihalek I, Lichtarge O, Bertuch AA (2007) Distinct faces of the Ku heterodimer mediate DNA repair and telomeric functions. *Nature Publishing Group* 14:301–307.
141. Song K, Jung D, Jung Y, Lee SG, Lee I (2000) Interaction of human Ku70 with TRF2. *FEBS Lett* 481:81–85.
142. Celli GB, Denchi EL, de Lange T (2006) Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination. *Nat Cell Biol* 8:885–890.
143. Downs JA, Jackson SP (2004) A means to a DNA end: the many roles of Ku. *Nat Rev Mol Cell Biol* 5:367–378.
144. Teoh G et al. (1998) The 86-kD subunit of Ku autoantigen mediates homotypic and heterotypic adhesion of multiple myeloma cells. *J Clin Invest* 101:1379–1388.
145. Lynch EM, Moreland RB, Ginis I, Perrine SP, Faller DV (2001) Hypoxia-activated ligand HAL-1/13 is lupus autoantigen Ku80 and mediates lymphoid cell adhesion in vitro. *Am. J. Physiol., Cell Physiol.* 280:C897–911.
146. Zheng Y, Ao Z, Wang B, Jayappa KD, Yao X (2011) Host protein Ku70 binds and protects HIV-1 integrase from proteasomal degradation and is required for HIV replication. *J Biol Chem* 286:17722–17735.
147. Munakata Y et al. (2005) Ku80 autoantigen as a cellular coreceptor for human parvovirus B19 infection. *Blood* 106:3449–3456.
148. Martinez JJ, Seveau S, Veiga E, Matsuyama S, Cossart P (2005) Ku70, a component of DNA-dependent protein kinase, is a mammalian receptor for *Rickettsia conorii*. *Cell* 123:1013–1023.
149. Monferran S, Muller C, Mourey L, Frit P, Salles B (2004) The Membrane-associated form of the DNA repair protein Ku is involved in cell adhesion to fibronectin. *J Mol Biol* 337:503–511.
150. Monferran S, Paupert J, Dauvillier S, Salles B, Muller C (2004) The membrane form of the DNA repair protein Ku interacts at the cell surface with metalloproteinase 9. *EMBO J* 23:3758–3768.
151. Paupert J, Dauvillier S, Salles B, Muller C (2007) Transport of the leaderless protein Ku on the cell surface of activated monocytes regulates their migratory abilities. *EMBO Rep* 8:583–588.

152. Sawada M et al. (2003) Ku70 suppresses the apoptotic translocation of Bax to mitochondria. *Nat Cell Biol* 5:320–329.
153. Cohen HY et al. (2004) Calorie restriction promotes mammalian cell survival by inducing the SIRT1 deacetylase. *Science* 305:390–392.
154. Shuman S, Glickman MS (2007) Bacterial DNA repair by non-homologous end joining. *Nat Rev Microbiol* 5:852–861.
155. Weller GR et al. (2002) Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* 297:1686–1689.
156. Dmitrieva NI, Chen H-T, Nussenzweig A, Burg MB (2009) Knockout of Ku86 accelerates cellular senescence induced by high NaCl. *Aging (Albany NY)* 1:245–253.
157. Riha K, Shippen DE (2003) Ku is required for telomeric C-rich strand maintenance but not for end-to-end chromosome fusions in Arabidopsis. *Proc Natl Acad Sci USA* 100:611–615.
158. Adachi N, Ishino T, Ishii Y, Takeda S, Koyama H (2001) DNA ligase IV-deficient cells are more resistant to ionizing radiation in the absence of Ku70: Implications for DNA double-strand break repair. *Proc Natl Acad Sci USA* 98:12109–12113.
159. Wang H et al. (2001) Efficient rejoining of radiation-induced DNA double-strand breaks in vertebrate cells deficient in genes of the RAD52 epistasis group. *Oncogene* 20:2212–2224.
160. Gu Y, Jin S, Gao Y, Weaver DT, Alt FW (1997) Ku70-deficient embryonic stem cells have increased ionizing radiosensitivity, defective DNA end-binding activity, and inability to support V(D)J recombination. *Proc Natl Acad Sci USA* 94:8076–8081.
161. Nussenzweig A et al. (1996) Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature* 382:551–555.
162. Gu Y et al. (1997) Growth retardation and leaky SCID phenotype of Ku70-deficient mice. *Immunity* 7:653–665.
163. Blackburn EH (2001) Switching and signaling at the telomere. *Cell* 106:661–673.
164. de Lange T (2002) Protection of mammalian telomeres. *Oncogene* 21:532–540.
165. Donate LE, Blasco MA (2011) Telomeres in cancer and ageing. *Philos Trans R Soc Lond, B, Biol Sci* 366:76–84.
166. Moshous D et al. (2000) A new gene involved in DNA double-strand break repair and V(D)J recombination is located on human chromosome 10p. *Hum Mol Genet* 9:583–588.
167. Lewis SM, Wu GE (1997) The origins of V(D)J recombination. *Cell* 88:159–162.
168. DeFazio LG, Stansel RM, Griffith JD, Chu G (2002) Synapsis of DNA ends by DNA-dependent protein kinase. *EMBO J* 21:3192–3200.
169. Boskovic J et al. (2003) Visualization of DNA-induced conformational

- changes in the DNA repair kinase DNA-PKcs. *EMBO J* 22:5875–5882.
170. Yoo S, Dynan WS (1999) Geometry of a complex formed by double strand break repair proteins at a single DNA end: recruitment of DNA-PKcs induces inward translocation of Ku protein. *Nucleic Acids Res* 27:4679–4686.
171. West RB, Yaneva M, Lieber MR (1998) Productive and nonproductive complexes of Ku and DNA-dependent protein kinase at DNA termini. *Mol Cell Biol* 18:5908–5920.
172. Chan DW et al. (2002) Autophosphorylation of the DNA-dependent protein kinase catalytic subunit is required for rejoining of DNA double-strand breaks. *Genes Dev* 16:2333–2338.
173. Douglas P et al. (2002) Identification of in vitro and in vivo phosphorylation sites in the catalytic subunit of the DNA-dependent protein kinase. *Biochem J* 368:243–251.
174. Meek K, Douglas P, Cui X, Ding Q, Lees-Miller SP (2007) trans Autophosphorylation at DNA-dependent protein kinase's two major autophosphorylation site clusters facilitates end processing but not end joining. *Mol Cell Biol* 27:3881–3890.
175. Hammel M et al. (2010) Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *J Biol Chem* 285:1414–1423.
176. Dobbs TA et al. (2010) A structural model for regulation of NHEJ by DNA-PKcs autophosphorylation. *DNA Repair (Amst)* 9:1307–1314.
177. Uematsu N et al. (2007) Autophosphorylation of DNA-PKCS regulates its dynamics at DNA double-strand breaks. *J Cell Biol* 177:219–229.
178. Chan DW, Ye R, Veillette CJ, Lees-Miller SP (1999) DNA-dependent protein kinase phosphorylation sites in Ku 70/80 heterodimer. *Biochemistry* 38:1819–1828.
179. Lee K-J, Jovanovic M, Udayakumar D, Bladen CL, Dynan WS (2004) Identification of DNA-PKcs phosphorylation sites in XRCC4 and effects of mutations at these sites on DNA end joining in a cell-free system. *DNA Repair (Amst)* 3:267–276.
180. Yu Y et al. (2008) DNA-PK and ATM phosphorylation sites in XLF/Cernunnos are not required for repair of DNA double strand breaks. *DNA Repair (Amst)* 7:1680–1692.
181. Ma Y et al. (2005) The DNA-dependent protein kinase catalytic subunit phosphorylation sites in human Artemis. *J Biol Chem* 280:33839–33846.
182. Ma Y, Pannicke U, Schwarz K, Lieber MR (2002) Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* 108:781–794.
183. Zhang X et al. (2004) Artemis is a phosphorylation target of ATM and ATR and is involved in the G2/M DNA damage checkpoint response. *Mol Cell Biol* 24:9207–9220.



184. Poinsignon C et al. (2004) Phosphorylation of Artemis following irradiation-induced DNA damage. *Eur J Immunol* 34:3146–3155.
185. Chen L et al. (2005) Ataxia-telangiectasia-mutated dependent phosphorylation of Artemis in response to DNA damage. *Cancer Sci* 96:134–141.
186. Riballo E et al. (2004) A pathway of double-strand break rejoining dependent upon ATM, Artemis, and proteins locating to gamma-H2AX foci. *Mol Cell* 16:715–724.
187. Ropars V et al. (2011) Structural characterization of filaments formed by human Xrcc4-Cernunnos/XLF complex involved in nonhomologous DNA end-joining. *Proceedings of the National Academy of Sciences* 108:12663–12668.
188. Sibanda BL et al. (2001) Crystal structure of an Xrcc4-DNA ligase IV complex. *Nat Struct Biol* 8:1015–1019.
189. Callebaut I et al. (2006) Cernunnos interacts with the XRCC4 x DNA-ligase IV complex and is homologous to the yeast nonhomologous end-joining factor Nej1. *J Biol Chem* 281:13857–13860.
190. Ellenberger T, Tomkinson AE (2008) Eukaryotic DNA ligases: structural and functional insights. *Annu Rev Biochem* 77:313–338.
191. Caldecott K (2007) Mammalian single-strand break repair: Mechanisms and links with chromatin. *DNA Repair (Amst)* 6:443–453.
192. Bentley DJ et al. (2002) DNA ligase I null mouse cells show normal DNA repair activity but altered DNA replication and reduced genome stability. *J. Cell. Sci.* 115:1551–1561.
193. Goetz JD-M, Motycka TA, Han M, Jasin M, Tomkinson AE (2005) Reduced repair of DNA double-strand breaks by homologous recombination in a DNA ligase I-deficient human cell line. *DNA Repair (Amst)* 4:649–654.
194. Wang H et al. (2005) DNA ligase III as a candidate component of backup pathways of nonhomologous end joining. *Cancer Res* 65:4020–4030.
195. Wei YF et al. (1995) Molecular cloning and expression of human cDNAs encoding a novel DNA ligase IV and DNA ligase III, an enzyme active in DNA repair and recombination. *Mol Cell Biol* 15:3206–3216.
196. Wu P-Y et al. (2009) Structural and functional interaction between the human DNA repair proteins DNA ligase IV and XRCC4. *Mol Cell Biol* 29:3163–3172.
197. Costantini S, Woodbine L, Andreoli L, Jeggo PA, Vindigni A (2007) Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK. *DNA Repair (Amst)* 6:712–722.
198. Li Z et al. (1995) The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. *Cell* 83:1079–1089.
199. Bryans M, Valenzano MC, Stamato TD (1999) Absence of DNA ligase IV protein in XR-1 cells: evidence for stabilization by XRCC4. *Mutat Res*

- 433:53–58.
200. Hammel M, Yu Y, Fang S, Lees-Miller SP, Tainer JA (2010) XLF regulates filament architecture of the XRCC4-ligase IV complex. *Structure* 18:1431–1442.
  201. Wang Y-G, Nnakwe C, Lane WS, Modesti M, Frank KM (2004) Phosphorylation and regulation of DNA ligase IV stability by DNA-dependent protein kinase. *J Biol Chem* 279:37282–37290.
  202. Leber R, Wise TW, Mizuta R, Meek K (1998) The XRCC4 gene product is a target for and interacts with the DNA-dependent protein kinase. *J Biol Chem* 273:1794–1801.
  203. Buck D et al. (2006) Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell* 124:287–299.
  204. Ahnesorg P, Smith P, Jackson SP (2006) XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell* 124:301–313.
  205. Gu J, Lu H, Tsai AG, Schwarz K, Lieber MR (2007) Single-stranded DNA ligation and XLF-stimulated incompatible DNA end ligation by the XRCC4-DNA ligase IV complex: influence of terminal DNA sequence. *Nucleic Acids Res* 35:5755–5762.
  206. Tsai CJ, Kim SA, Chu G (2007) Cernunnos/XLF promotes the ligation of mismatched and noncohesive DNA ends. *Proc Natl Acad Sci USA* 104:7851–7856.
  207. Grawunder U, Zimmer D, Fugmann S, Schwarz K, Lieber MR (1998) DNA ligase IV is essential for V(D)J recombination and DNA double-strand break repair in human precursor lymphocytes. *Mol Cell* 2:477–484.
  208. Gatz SA et al. (2011) Requirement for DNA Ligase IV during Embryonic Neuronal Development. *Journal of Neuroscience* 31:10088–10100.
  209. Barnes DE, Stamp G, Rosewell I, Denzel A, Lindahl T (1998) Targeted disruption of the gene encoding DNA ligase IV leads to lethality in embryonic mice. *Curr Biol* 8:1395–1398.
  210. Frank KM et al. (2000) DNA ligase IV deficiency in mice leads to defective neurogenesis and embryonic lethality via the p53 pathway. *Mol Cell* 5:993–1002.
  211. Gao Y et al. (1998) A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* 95:891–902.
  212. Li G et al. (2008) Lymphocyte-specific compensation for XLF/cernunnos end-joining functions in V(D)J recombination. *Mol Cell* 31:631–640.
  213. Moon AF et al. (2007) The X family portrait: structural insights into biological functions of X family polymerases. *DNA Repair (Amst)* 6:1709–1725.
  214. Ramadan K, Shevelev I, Hübscher U (2004) Opinion: The DNA-polymerase-X family: controllers of DNA quality? *Nat Rev Mol Cell Biol*

- 5:1038–1043.
215. Chang LM, Bolland FJ (1986) Molecular biology of terminal transferase. *CRC Crit. Rev. Biochem.* 21:27–52.
216. Yano K-I et al. (2008) Ku recruits XLF to DNA double-strand breaks. *EMBO Rep* 9:91–96.
217. Tomimatsu N et al. (2007) Ku70/80 modulates ATM and ATR signaling pathways in response to DNA double strand breaks. *J Biol Chem* 282:10138–10145.
218. Rivera-Calzada A et al. (2005) Three-dimensional structure and regulation of the DNA-dependent protein kinase catalytic subunit (DNA-PKcs). *Structure* 13:243–255.
219. Koch CA et al. (2004) Xrcc4 physically links DNA end processing by polynucleotide kinase to DNA ligation by DNA ligase IV. *EMBO J* 23:3874–3885.
220. Mani RS et al. (2010) Dual modes of interaction between XRCC4 and polynucleotide kinase/phosphatase: implications for nonhomologous end joining. *J Biol Chem* 285:37619–37629.
221. Williams RS et al. (2005) Structural basis for phosphorylation-dependent signaling in the DNA-damage response. *Biochem. Cell Biol.* 83:721–727.
- 222.
223. Kysela B et al. (2003) Ku stimulation of DNA ligase IV-dependent ligation requires inward movement along the DNA molecule. *J Biol Chem* 278:22466–22474.
- 224.
225. Postow L et al. (2008) Ku80 removal from DNA through double strand break-induced ubiquitylation. *J Cell Biol* 182:467–479.
226. Budman J, Kim SA, Chu G (2007) Processing of DNA for nonhomologous end-joining is controlled by kinase activity and XRCC4/ligase IV. *J Biol Chem* 282:11950–11959.
227. Zhang Y et al. (2007) Role of Dnl4-Lif1 in nonhomologous end-joining repair complex assembly and suppression of homologous recombination. *Nature Publishing Group* 14:639–646.
228. Simsek D, Jasin M (2010) Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4-ligase IV during chromosomal translocation formation. *Nat Struct Mol Biol* 17:410–416.
229. Ball LG, Xiao W (2005) Molecular basis of ataxia telangiectasia and related diseases. *Acta Pharmacol. Sin.* 26:897–907.
230. O'Driscoll M, Gennery AR, Seidel J, Concannon P, Jeggo PA (2004) An overview of three new disorders associated with genetic instability: LIG4 syndrome, RS-SCID and ATR-Seckel syndrome. *DNA Repair (Amst)* 3:1227–1235.
231. Seemanová E et al. (2006) Nijmegen breakage syndrome (NBS) with neurological abnormalities and without chromosomal instability. *J. Med.*

- Genet.* 43:218–224.
232. Varon R et al. (1998) Nibrin, a novel DNA double-strand break repair protein, is mutated in Nijmegen breakage syndrome. *Cell* 93:467–476.
233. Digweed M, Sperling K (2004) Nijmegen breakage syndrome: clinical manifestation of defective response to DNA double-strand breaks. *DNA Repair (Amst)* 3:1207–1217.
234. Stewart GS et al. (1999) The DNA double-strand break repair gene hMRE11 is mutated in individuals with an ataxia-telangiectasia-like disorder. *Cell* 99:577–587.
235. Taylor AMR, Groom A, Byrd PJ (2004) Ataxia-telangiectasia-like disorder (ATLD)-its clinical presentation and molecular basis. *DNA Repair (Amst)* 3:1219–1225.
236. van der Burg M, van Dongen JJ, van Gent DC (2009) DNA-PKcs deficiency in human: long predicted, finally found. *Curr Opin Allergy Clin Immunol* 9:503–509.
237. Moshous D et al. (2001) Artemis, a novel DNA double-strand break repair/V(D)J recombination protein, is mutated in human severe combined immune deficiency. *Cell* 105:177–186.
238. Riballo E et al. (1999) Identification of a defect in DNA ligase IV in a radiosensitive leukaemia patient. *Curr Biol* 9:699–702.
239. O'Driscoll M et al. (2001) DNA ligase IV mutations identified in patients exhibiting developmental delay and immunodeficiency. *Mol Cell* 8:1175–1185.
240. Ben-Omran TI, Cerosaletti K, Concannon P, Weitzman S, Nezarati MM (2005) A patient with mutations in DNA Ligase IV: clinical features and overlap with Nijmegen breakage syndrome. *Am. J. Med. Genet. A* 137A:283–287.
241. Girard P-M, Kysela B, Härer CJ, Doherty AJ, Jeggo PA (2004) Analysis of DNA ligase IV mutations found in LIG4 syndrome patients: the impact of two linked polymorphisms. *Hum Mol Genet* 13:2369–2376.
242. Chistiakov DA (2010) Ligase IV syndrome. *Adv. Exp. Med. Biol.* 685:175–185.
243. van der Burg M et al. (2006) A new type of radiosensitive T-B-NK+ severe combined immunodeficiency caused by a LIG4 mutation. *J Clin Invest* 116:137–145.
244. Wang H et al. (2001) Genetic evidence for the involvement of DNA ligase IV in the DNA-PK-dependent pathway of non-homologous end joining in mammalian cells. *Nucleic Acids Res* 29:1653–1660.
245. Nash RA, Caldecott KW, Barnes DE, Lindahl T (1997) XRCC1 protein interacts with one of two distinct forms of DNA ligase III. *Biochemistry* 36:5207–5211.
246. Mackey ZB et al. (1997) An alternative splicing event which occurs in mouse pachytene spermatocytes generates a form of DNA ligase III with

- distinct biochemical properties that may function in meiotic recombination. *Mol Cell Biol* 17:989–998.
247. Caldecott KW, McKeown CK, Tucker JD, Ljungquist S, Thompson LH (1994) An interaction between the mammalian DNA repair protein XRCC1 and DNA ligase III. *Mol Cell Biol* 14:68–76.
248. Audebert M, Salles B, Calsou P (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* 279:55117–55126.
249. Hohegger H et al. (2006) Parp-1 protects homologous recombination from interference by Ku and Ligase IV in vertebrate cells. *EMBO J* 25:1305–1314.
250. Rosidi B et al. (2008) Histone H1 functions as a stimulatory factor in backup pathways of NHEJ. *Nucleic Acids Res* 36:1610–1623.
251. Corneo B et al. (2007) Rag mutations reveal robust alternative end joining. *Nature* 449:483–486.
252. Deriano L, Stracker TH, Baker A, Petrini JHJ, Roth DB (2009) Roles for NBS1 in alternative nonhomologous end-joining of V(D)J recombination intermediates. *Mol Cell* 34:13–25.
253. De A, Campbell C (2007) A novel interaction between DNA ligase III and DNA polymerase gamma plays an essential role in mitochondrial DNA stability. *Biochem J* 402:175–186.
254. Fattah F et al. (2010) Ku regulates the non-homologous end joining pathway choice of DNA double-strand break repair in human somatic cells. *PLoS Genet* 6:e1000855.
255. Bahmed K, Seth A, Nitiss KC, Nitiss JL (2010) End-processing during non-homologous end-joining: a role for exonuclease 1. *Nucleic Acids Res*.
256. Della-Maria J et al. (2011) hMre11/hRad50/Nbs1 and DNA ligase III /XRCC1 act together in an alternative non-homologous end joining pathway. *J Biol Chem*.
257. Rahal EA et al. (2010) ATM regulates Mre11-dependent DNA end-degradation and microhomology-mediated end joining. *Cell Cycle* 9:2866–2877.
258. Rass E et al. (2009) Role of Mre11 in chromosomal nonhomologous end joining in mammalian cells. *Nat Struct Mol Biol* 16:819–824.
259. Xie A, Kwok A, Scully R (2009) Role of mammalian Mre11 in classical and alternative nonhomologous end joining. *Nature Publishing Group* 16:814–818.
260. Rai R et al. (2010) The function of classical and alternative non-homologous end-joining pathways in the fusion of dysfunctional telomeres. *EMBO J* 29:2598–2610.
261. Perez-Jannotti RM, Klein SM, Bogenhagen DF (2001) Two forms of mitochondrial DNA ligase III are produced in *Xenopus laevis* oocytes. *J Biol Chem* 276:48978–48987.

262. Lakshmipathy U, Campbell C (1999) The human DNA ligase III gene encodes nuclear and mitochondrial proteins. *Mol Cell Biol* 19:3869–3876.
263. Dulic A et al. (2001) BRCT domain interactions in the heterodimeric DNA repair protein XRCC1-DNA ligase III. *Biochemistry* 40:5906–5913.
264. Caldecott KW, Tucker JD, Stanker LH, Thompson LH (1995) Characterization of the XRCC1-DNA ligase III complex in vitro and its absence from mutant hamster cells. *Nucleic Acids Res* 23:4836–4843.
265. Tebbs RS et al. (1999) Requirement for the Xrcc1 DNA base excision repair gene during early mouse development. *Dev Biol* 208:513–529.
266. Puebla-Osorio N, Lacey DB, Alt FW, Zhu C (2006) Early embryonic lethality due to targeted inactivation of DNA ligase III. *Mol Cell Biol* 26:3935–3941.
267. Lakshmipathy U, Campbell C (2000) Mitochondrial DNA ligase III function is independent of Xrcc1. *Nucleic Acids Res* 28:3880–3886.
268. Lakshmipathy U, Campbell C (2001) Antisense-mediated decrease in DNA ligase III expression results in reduced mitochondrial DNA integrity. *Nucleic Acids Res* 29:668–676.
269. Simsek D et al. (2011) Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. *Nature* 471:245–248.
270. Nair PA et al. (2007) Structural basis for nick recognition by a minimal pluripotent DNA ligase. *Nature Publishing Group* 14:770–778.
271. Ruhanen H et al. (2011) Involvement of DNA ligase III and ribonuclease H1 in mitochondrial DNA replication in cultured human cells. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*.
272. Simsek D et al. (2011) DNA Ligase III Promotes Alternative Nonhomologous End-Joining during Chromosomal Translocation Formation. *PLoS Genet* 7:e1002080.
273. Mani R-S, Chinnaiyan AM (2010) Triggers for genomic rearrangements: insights into genomic, cellular and environmental influences. *Nat Rev Genet* 11:819–829.
274. McVey M, Lee SE (2008) MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet* 24:529–538.
275. Guirouilh-Barbat J, Rass E, Plo I, Bertrand P, Lopez BS (2007) Defects in XRCC4 and KU80 differentially affect the joining of distal nonhomologous ends. *Proc Natl Acad Sci USA* 104:20902–20907.
276. Fattah FJ, Lichter NF, Fattah KR, Oh S, Hendrickson EA (2008) Ku70, an essential gene, modulates the frequency of rAAV-mediated gene targeting in human somatic cells. *Proc Natl Acad Sci USA* 105:8703–8708.
277. Iizumi S et al. (2008) Impact of non-homologous end-joining deficiency on random and targeted DNA integration: implications for gene targeting. *Nucleic Acids Res* 36:6333–6342.
278. Vasileva A, Linden RM, Jessberger R (2006) Homologous recombination is required for AAV-mediated gene targeting. *Nucleic Acids Res* 34:3345–

- 3360.
279. Budman J, Chu G (2005) Processing of DNA for nonhomologous end-joining by cell-free extract. *EMBO J* 24:849–860.
280. Bothmer A et al. (2010) 53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination. *J Exp Med* 207:855–865.
281. Langerak P, Mejia-Ramirez E, Limbo O, Russell P (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:e1002271.
282. Mimitou EP, Symington LS (2009) DNA end resection: many nucleases make light work. *DNA Repair (Amst)* 8:983–995.
283. Helmink BA et al. (2010) H2AX prevents CtIP-mediated DNA end resection and aberrant repair in G1-phase lymphocytes. *Nature*.
284. Yun MH, Hiom K (2009) CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. *Nature* 459:460–463.
285. Wu W, Wang M, Mussfeldt T, Iliakis G (2008) Enhanced use of backup pathways of NHEJ in G2 in Chinese hamster mutant cells with defects in the classical pathway of NHEJ. *Radiat Res* 170:512–520.
286. Iliakis G (2009) Backup pathways of NHEJ in cells of higher eukaryotes: cell cycle dependence. *Radiother Oncol* 92:310–315.
287. Wu W et al. (2008) Repair of radiation induced DNA double strand breaks by backup NHEJ is enhanced in G2. *DNA Repair (Amst)* 7:329–338.
288. Yu X, Baer R (2000) Nuclear localization and cell cycle-specific expression of CtIP, a protein that associates with the BRCA1 tumor suppressor. *J Biol Chem* 275:18541–18549.
289. Yu X, Chen J (2004) DNA damage-induced cell cycle checkpoint control requires CtIP, a phosphorylation-dependent binding partner of BRCA1 C-terminal domains. *Mol Cell Biol* 24:9478–9486.
290. Chen L, Nievera CJ, Lee AY-L, Wu X (2008) Cell cycle-dependent complex formation of BRCA1.CtIP.MRN is important for DNA double-strand break repair. *J Biol Chem* 283:7713–7720.
291. Yu X, Wu LC, Bowcock AM, Aronheim A, Baer R (1998) The C-terminal (BRCT) domains of BRCA1 interact in vivo with CtIP, a protein implicated in the CtBP pathway of transcriptional repression. *J Biol Chem* 273:25388–25392.
292. Watson JD (1972) Origin of concatemeric T7 DNA. *Nature New Biol.* 239:197–201.
293. Olovnikov AM (1973) A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *Journal of Theoretical Biology* 41:181–190.
294. Heiss NS et al. (1998) X-linked dyskeratosis congenita is caused by

- mutations in a highly conserved gene with putative nucleolar functions. *Nat Genet* 19:32–38.
295. Knight SW et al. (1999) X-linked dyskeratosis congenita is predominantly caused by missense mutations in the DKC1 gene. *Am. J. Hum. Genet.* 65:50–58.
296. Bodnar AG et al. (1998) Extension of life-span by introduction of telomerase into normal human cells. *Science* 279:349–352.
297. Box C, Rogers SJ, Mendiola M, Eccles SA (2010) Tumour-microenvironmental interactions: paths to progression and targets for treatment. *Semin. Cancer Biol.* 20:128–138.
298. Blasco MA (2005) Telomeres and human disease: ageing, cancer and beyond. *Nat Rev Genet* 6:611–622.
299. Nabetani A, Ishikawa F (2011) Alternative lengthening of telomeres pathway: recombination-mediated telomere maintenance mechanism in human cells. *J. Biochem.* 149:5–14.
300. Cesare AJ, Reddel RR (2010) Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet* 11:319–330.
301. Samassekou O, Gadjji M, Drouin R, Yan J (2010) Sizing the ends: normal length of human telomeres. *Ann. Anat.* 192:284–291.
302. Slijepcevic P (2006) The role of DNA damage response proteins at telomeres--an “integrative” model. *DNA Repair (Amst)* 5:1299–1306.
303. Bai Y, Murnane JP (2003) Telomere instability in a human tumor cell line expressing NBS1 with mutations at sites phosphorylated by ATM. *Mol Cancer Res* 1:1058–1069.
304. Bae NS, Baumann P (2007) A RAP1/TRF2 complex inhibits nonhomologous end-joining at human telomeric DNA ends. *Mol Cell* 26:323–334.
305. Bombarde O et al. (2010) TRF2/RAP1 and DNA-PK mediate a double protection against joining at telomeric ends. *EMBO J* 29:1573–1584.
306. Wang RC, Smogorzewska A, de Lange T (2004) Homologous recombination generates T-loop-sized deletions at human telomeres. *Cell* 119:355–368.
307. Griffith JD et al. (1999) Mammalian telomeres end in a large duplex loop. *Cell* 97:503–514.
308. Hockemeyer D et al. (2007) Telomere protection by mammalian Pot1 requires interaction with Tpp1. *Nature Publishing Group* 14:754–761.
309. Guo X et al. (2007) Dysfunctional telomeres activate an ATM-ATR-dependent DNA damage response to suppress tumorigenesis. *EMBO J* 26:4709–4719.
310. Martinez P et al. (2010) Mammalian Rap1 controls telomere function and gene expression through binding to telomeric and extratelomeric sites. *Nat Cell Biol* 12:768–780.
311. Celli GB, de Lange T (2005) DNA processing is not required for ATM-mediated telomere damage response after TRF2 deletion. *Nat Cell Biol*



- 7:712–718.
312. Karlseder J, Broccoli D, Dai Y, Hardy S, de Lange T (1999) p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. *Science* 283:1321–1325.
  313. Denchi EL, de Lange T (2007) Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1. *Nature* 448:1068–1071.
  314. Barrientos KS et al. (2008) Distinct functions of POT1 at telomeres. *Mol Cell Biol* 28:5251–5264.
  315. Ranganathan V et al. (2001) Rescue of a telomere length defect of Nijmegen breakage syndrome cells requires NBS and telomerase catalytic subunit. *Curr Biol* 11:962–966.
  316. Misri S, Pandita S, Kumar R, Pandita TK (2008) Telomeres, histone code, and DNA damage response. *Cytogenet. Genome Res.* 122:297–307.
  317. Denchi EL (2009) Give me a break: how telomeres suppress the DNA damage response. *DNA Repair (Amst)* 8:1118–1126.
  318. Smogorzewska A, de Lange T (2002) Different telomere damage signaling pathways in human and mouse cells. *EMBO J* 21:4338–4348.
  319. d'Adda di Fagagna F et al. (2003) A DNA damage checkpoint response in telomere-initiated senescence. *Nature* 426:194–198.
  320. Deng Y, Guo X, Ferguson DO, Chang S (2009) Multiple roles for MRE11 at uncapped telomeres. *Nature* 460:914–918.
  321. Yang D, He Q, Kim H, Ma W, Songyang Z (2011) TIN2 protein dyskeratosis congenita missense mutants are defective in association with telomerase. *J Biol Chem* 286:23022–23030.
  322. Zijlmans JM et al. (1997) Telomeres in the mouse have large inter-chromosomal variations in the number of T2AG3 repeats. *Proc Natl Acad Sci USA* 94:7423–7428.
  323. Palm W, Hockemeyer D, Kibe T, de Lange T (2009) Functional dissection of human and mouse POT1 proteins. *Mol Cell Biol* 29:471–482.
  324. Baum C et al. (2003) Side effects of retroviral gene transfer into hematopoietic stem cells. *Blood* 101:2099–2114.
  325. Talbot D et al. (1989) A dominant control region from the human beta-globin locus conferring integration site-independent gene expression. *Nature* 338:352–355.
  326. Hacein-Bey-Abina S et al. (2003) A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N. Engl. J. Med.* 348:255–256.
  327. Vasileva A, Jessberger R (2005) Precise hit: adeno-associated virus in gene targeting. *Nat Rev Microbiol* 3:837–847.
  328. Sonoda E, Takata M, Yamashita YM, Morrison C, Takeda S (2001) Homologous DNA recombination in vertebrate cells. *Proc Natl Acad Sci USA* 98:8388–8394.
  329. Russell DW, Hirata RK (1998) Human gene targeting by viral vectors. *Nat*

- Genet* 18:325–330.
330. Khan IF et al. (2011) AAV-mediated gene targeting methods for human cells. *Nat Protoc* 6:482–501.
  331. Büning H, Perabo L, Coutelle O, Quadt-Humme S, Hallek M (2008) Recent developments in adeno-associated virus vector technology. *J Gene Med* 10:717–733.
  332. Schultz BR, Chamberlain JS (2008) Recombinant adeno-associated virus transduction and integration. *Mol Ther* 16:1189–1199.
  333. Chamberlain JR et al. (2008) Gene targeting of mutant COL1A2 alleles in mesenchymal stem cells from individuals with osteogenesis imperfecta. *Mol Ther* 16:187–193.
  334. Kohli M, Rago C, Lengauer C, Kinzler KW, Vogelstein B (2004) Facile methods for generating human somatic cell gene knockouts using recombinant adeno-associated viruses. *Nucleic Acids Res* 32:e3.
  335. Topaloglu O, Hurley PJ, Yildirim O, Civin CI, Bunz F (2005) Improved methods for the generation of human gene knockout and knockin cell lines. *Nucleic Acids Res* 33:e158.
  336. Kim J-S, Bonifant C, Bunz F, Lane WS, Waldman T (2008) Epitope tagging of endogenous genes in diverse human cell lines. *Nucleic Acids Res* 36:e127.
  337. Zhang X et al. (2008) Epitope tagging of endogenous proteins for genome-wide ChIP-chip studies. *Nat Methods* 5:163–165.
  338. Bennett CB, Lewis AL, Baldwin KK, Resnick MA (1993) Lethality induced by a single site-specific double-strand break in a dispensable yeast plasmid. *Proc Natl Acad Sci USA* 90:5613–5617.
  339. Kass EM, Jasin M (2010) Collaboration and competition between DNA double-strand break repair pathways. *FEBS Lett* 584:3703–3708.
  340. Hartlerode AJ, Scully R (2009) Mechanisms of double-strand break repair in somatic mammalian cells. *Biochem J* 423:157–168.
  341. Nussenzweig A, Nussenzweig MC (2007) A backup DNA repair pathway moves to the forefront. *Cell* 131:223–225.
  342. Mladenov E, Iliakis G (2011) Induction and repair of DNA double strand breaks: the increasing spectrum of non-homologous end joining pathways. *Mutat Res* 711:61–72.
  343. Tichy ED et al. (2010) Mouse embryonic stem cells, but not somatic cells, predominantly use homologous recombination to repair double-strand DNA breaks. *Stem Cells Dev.* 19:1699–1711.
  344. Ayora S et al. (2011) Double-strand break repair in bacteria: a view from *Bacillus subtilis*. *FEMS Microbiology Reviews*:no–no.
  345. Heyer W-D, Ehmsen KT, Liu J (2010) Regulation of Homologous Recombination in Eukaryotes. *Annu Rev Genet* 44:113–139.
  346. Lieber MR, Yu K, Raghavan SC (2006) Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. *DNA Repair (Amst)* 5:1234–1245.

347. Han L, Yu K (2008) Altered kinetics of nonhomologous end joining and class switch recombination in ligase IV-deficient B cells. *J Exp Med* 205:2745–2753.
348. Yan CT et al. (2007) IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* 449:478–482.
349. Soulas-Sprauel P et al. (2007) Role for DNA repair factor XRCC4 in immunoglobulin class switch recombination. *J Exp Med* 204:1717–1727.
350. de Lange T (2004) T-loops and the origin of telomeres. *Nat Rev Mol Cell Biol* 5:323–329.
351. Neal JA, Meek K (2011) Choosing the right path: Does DNA-PK help make the decision? *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 711:73–86.
352. Soubeyrand S et al. (2006) Artemis phosphorylated by DNA-dependent protein kinase associates preferentially with discrete regions of chromatin. *J Mol Biol* 358:1200–1211.
353. Grawunder U et al. (1997) Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* 388:492–495.
354. Chen L, Trujillo K, Sung P, Tomkinson AE (2000) Interactions of the DNA ligase IV-XRCC4 complex with DNA ends and the DNA-dependent protein kinase. *J Biol Chem* 275:26196–26205.
355. Riballo E et al. (2009) XLF-Cernunnos promotes DNA ligase IV-XRCC4 re-adenylation following ligation. *Nucleic Acids Res* 37:482–492.
356. Frank KM et al. (1998) Late embryonic lethality and impaired V(D)J recombination in mice lacking DNA ligase IV. *Nature* 396:173–177.
357. Puebla-Osorio N, Zhu C (2008) DNA damage and repair during lymphoid development: antigen receptor diversity, genomic integrity and lymphomagenesis. *Immunol. Res.* 41:103–122.
358. He DM, Lee SE, Hendrickson EA (1996) Restoration of X-ray and etoposide resistance, Ku-end binding activity and V(D) J recombination to the Chinese hamster xci-3 mutant by a hamster Ku86 cDNA. *Mutat Res* 363:43–56.
359. Take Y, Kumano M, Teraoka H, Nishimura S, Okuyama A (1996) DNA-dependent protein kinase inhibitor (OK-1035) suppresses p21 expression in HCT116 cells containing wild-type p53 induced by adriamycin. *Biochem Biophys Res Commun* 221:207–212.
360. Jiricny J (2006) The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* 7:335–346.
361. Eccleston J, Yan C, Yuan K, Alt FW, Selsing E (2011) Mismatch repair proteins MSH2, MLH1, and EXO1 are important for class-switch recombination events occurring in B cells that lack nonhomologous end joining. *J. Immunol.* 186:2336–2343.
362. Boland CR (1996) Roles of the DNA mismatch repair genes in colorectal tumorigenesis. *Int J Cancer* 69:47–49.
363. Ruis BL, Fattah KR, Hendrickson EA (2008) The catalytic subunit of DNA-

- dependent protein kinase regulates proliferation, telomere length, and genomic stability in human somatic cells. *Mol Cell Biol* 28:6182–6195.
364. Karanjawala ZE, Grawunder U, Hsieh CL, Lieber MR (1999) The nonhomologous DNA end joining pathway is important for chromosome stability in primary fibroblasts. *Curr Biol* 9:1501–1504.
365. Pacchierotti F, Ranaldi R (2006) Mechanisms and risk of chemically induced aneuploidy in mammalian germ cells. *Curr. Pharm. Des.* 12:1489–1504.
366. Fattah KR, Ruis BL, Hendrickson EA (2008) Mutations to Ku reveal differences in human somatic cell lines. *DNA Repair (Amst)* 7:762–774.
367. Chamberlain JR et al. (2004) Gene targeting in stem cells from individuals with osteogenesis imperfecta. *Science* 303:1198–1201.
368. Kim J-S, Lee C, Foxworth A, Waldman T (2004) B-Raf is dispensable for K-Ras-mediated oncogenesis in human cancer cells. *Cancer Res* 64:1932–1937.
369. Ninomiya Y, Suzuki K, Ishii C, Inoue H (2004) Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. *Proc Natl Acad Sci USA* 101:12248–12253.
370. Colot HV et al. (2006) A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. *Proc Natl Acad Sci USA* 103:10352–10357.
371. Sell C (2011) Ku circles the telomere? *Aging (Albany NY)* 3:344–345.
372. Massaad CA, Klann E (2011) Reactive oxygen species in the regulation of synaptic plasticity and memory. *Antioxid. Redox Signal.* 14:2013–2054.
373. Jeppesen DK, Bohr VA, Stevnsner T (2011) DNA repair deficiency in neurodegeneration. *Prog. Neurobiol.* 94:166–200.
374. Kontos HA (1989) Oxygen radicals in CNS damage. *Chem. Biol. Interact.* 72:229–255.
375. Kwon T, Huq E, Herrin DL (2010) Microhomology-mediated and nonhomologous repair of a double-strand break in the chloroplast genome of *Arabidopsis*. *Proc Natl Acad Sci USA* 107:13954–13959.
376. Cheng Q et al. (2011) Ku counteracts mobilization of PARP1 and MRN in chromatin damaged with DNA double-strand breaks. *Nucleic Acids Res.*
377. Samai P, Shuman S (2011) Functional dissection of the DNA interface of the nucleotidyltransferase domain of chloroella virus DNA ligase. *J Biol Chem* 286:13314–13326.
378. Nakai H et al. (2005) Large-scale molecular characterization of adeno-associated virus vector integration in mouse liver. *J Virol* 79:3606–3614.
379. Zhang Y, Jasin M (2010) An essential role for CtIP in chromosomal translocation formation through an alternative end-joining pathway. *Nat Struct Mol Biol.*
380. Brunet E et al. (2009) Chromosomal translocations induced at specified loci in human stem cells. *Proceedings of the National Academy of Sciences* 106:10620–10625.
381. Krishnakumar R, Kraus WL (2010) The PARP side of the nucleus:

- molecular actions, physiological outcomes, and clinical targets. *Mol Cell* 39:8–24.
382. Moser J et al. (2007) Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. *Mol Cell* 27:311–323.
383. Suganuma T, Workman JL (2011) Signals and combinatorial functions of histone modifications. *Annu Rev Biochem* 80:473–499.
384. Lundin C et al. (2005) Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks. *Nucleic Acids Res* 33:3799–3811.
385. Weterings E et al. (2009) The Ku80 carboxy terminus stimulates joining and artemis-mediated processing of DNA ends. *Mol Cell Biol* 29:1134–1142.
386. Hladnik U, Nyhan WL, Bertelli M (2008) Variable expression of HPRT deficiency in 5 members of a family with the same mutation. *Arch. Neurol.* 65:1240–1243.
387. Timson DJ, Singleton MR, Wigley DB (2000) DNA ligases in the repair and replication of DNA. *Mutat Res* 460:301–318.
388. Zheng L, Shen B (2011) Okazaki fragment maturation: nucleases take centre stage. *Journal of Molecular Cell Biology* 3:23–30.
389. Frosina G et al. (1996) Two pathways for base excision repair in mammalian cells. *J Biol Chem* 271:9573–9578.
390. Poyton RO, Ball KA, Castello PR (2009) Mitochondrial generation of free radicals and hypoxic signaling. *Trends Endocrinol. Metab.* 20:332–340.
391. Mookerjee SA, Divakaruni AS, Jastroch M, Brand MD (2010) Mitochondrial uncoupling and lifespan. *Mech. Ageing Dev.* 131:463–472.
392. Wiesner RJ, Zsurka G, Kunz WS (2006) Mitochondrial DNA damage and the aging process: facts and imaginations. *Free Radic. Res.* 40:1284–1294.
393. Kang D, Hamasaki N (2003) Mitochondrial oxidative stress and mitochondrial DNA. *Clin. Chem. Lab. Med.* 41:1281–1288.
394. Windhofer F, Wu W, Iliakis G (2007) Low levels of DNA ligases III and IV sufficient for effective NHEJ. *J Cell Physiol* 213:475–483.
395. Harrington J, Hsieh CL, Gerton J, Bosma G, Lieber MR (1992) Analysis of the defect in DNA end joining in the murine scid mutation. *Mol Cell Biol* 12:4758–4768.
396. Freitas AA, de Magalhaes JP (2011) A review and appraisal of the DNA damage theory of ageing. *Mutat Res* 728:12–22.
397. Ilnytsky Y, Kovalchuk O (2011) Non-targeted radiation effects-an epigenetic connection. *Mutat Res* 714:113–125.
398. Caldecott KW (2008) Single-strand break repair and disease. *Nat Rev Genet* 9:493–493.
399. Wojtyla A, Gladych M, Rubis B (2011) Human telomerase activity regulation. *Mol Biol Rep* 38:3339–3349.

400. Zhong Z-H et al. (2007) Disruption of telomere maintenance by depletion of the MRE11/RAD50/NBS1 complex in cells that use alternative lengthening of telomeres. *J Biol Chem* 282:29314–29322.
401. Kooistra R, Hooykaas PJJ, Steensma HY (2004) Efficient gene targeting in *Kluyveromyces lactis*. *Yeast* 21:781–792.
402. Yamana Y et al. (2005) Regulation of homologous integration in yeast by the DNA repair proteins Ku70 and RecQ. *Mol. Genet. Genomics* 273:167–176.
403. Goins CL, Gerik KJ, Lodge JK (2006) Improvements to gene deletion in the fungal pathogen *Cryptococcus neoformans*: absence of Ku proteins increases homologous recombination, and co-transformation of independent DNA molecules allows rapid complementation of deletion phenotypes. *Fungal Genet. Biol.* 43:531–544.
404. Pöggeler S, Kück U (2006) Highly efficient generation of signal transduction knockout mutants using a fungal strain deficient in the mammalian ku70 ortholog. *Gene* 378:1–10.
405. Meyer V et al. (2007) Highly efficient gene targeting in the *Aspergillus niger* kusA mutant. *J. Biotechnol.* 128:770–775.
406. Yusa K, Zhou L, Li MA, Bradley A, Craig NL (2011) A hyperactive piggyBac transposase for mammalian applications. *Proceedings of the National Academy of Sciences* 108:1531–1536.
407. Hirata RK, Russell DW (2000) Design and packaging of adeno-associated virus gene targeting vectors. *J Virol* 74:4612–4620.
408. Compton SA, Choi J-H, Cesare AJ, Ozgür S, Griffith JD (2007) Xrcc3 and Nbs1 are required for the production of extrachromosomal telomeric circles in human alternative lengthening of telomere cells. *Cancer Res* 67:1513–1519.
409. Watson JM, Shippen DE (2007) Telomere rapid deletion regulates telomere length in *Arabidopsis thaliana*. *Mol Cell Biol* 27:1706–1715.
410. Li B, Jog SP, Reddy S, Comai L (2008) WRN controls formation of extrachromosomal telomeric circles and is required for TRF2DeltaB-mediated telomere shortening. *Mol Cell Biol* 28:1892–1904.
411. Mimitou EP, Symington LS (2011) DNA end resection-Unraveling the tail. *DNA Repair (Amst)*.
412. Brugmans L, Kanaar R, Essers J (2007) Analysis of DNA double-strand break repair pathways in mice. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 614:95–108.
413. Karanjawala ZE et al. (2002) The embryonic lethality in DNA ligase IV-deficient mice is rescued by deletion of Ku: implications for unifying the heterogeneous phenotypes of NHEJ mutants. *DNA Repair (Amst)* 1:1017–1026.
414. Oh-McGinnis R, Jones MJ, Lefebvre L (2010) Applications of the site-specific recombinase Cre to the study of genomic imprinting. *Brief Funct Genomics* 9:281–293.

415. Haber JE (2006) Transpositions and translocations induced by site-specific double-strand breaks in budding yeast. *DNA Repair (Amst)* 5:998–1009.
416. Weinstock DM, Brunet E, Jasin M (2007) Formation of NHEJ-derived reciprocal chromosomal translocations does not require Ku70. *Nat Cell Biol* 9:978–981.
417. Weinstock DM, Brunet E, Jasin M (2008) Induction of Chromosomal Translocations in Mouse and Human Cells Using Site-Specific Endonucleases. *JNCI Monographs* 2008:20–24.
418. Artandi SE et al. (2000) Telomere dysfunction promotes non-reciprocal translocations and epithelial cancers in mice. *Nature* 406:641–645.
419. Baird DM (2010) Variation at the TERT locus and predisposition for cancer. *Expert Rev Mol Med* 12:e16.
420. Wang H et al. (2003) Biochemical evidence for Ku-independent backup pathways of NHEJ. *Nucleic Acids Res* 31:5377–5388.
421. Liang L et al. (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* 36:3297–3310.