ESSENTIAL ROLES FOR POLYMERASE THETA-MEDIATED END JOINING IN REPAIR OF CHROMOSOME BREAKS

David William Wyatt

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum in Genetics and Molecular Biology in the College of Arts and Sciences.

Chapel Hill
2016

Approved by:

Dale Ramsden

Cyrus Vaziri

Jeff Sekelsy

Shawn Ahmed

Mike Emanuele

ABSTRACT

David William Wyatt: Essential Roles for Polymerase Theta-Mediated End Joining in Repair of Chromosome Breaks
(Under the direction of Dale Ramsden)

DNA double strand breaks (DSBs) constitute a rare but lethal class of genomic damage that must be efficiently repaired. Deficiencies in DSB repair pathways manifest themselves as severe phenotypes including cancer predisposition, accelerated aging, and immunodeficiency. On the cellular level, failure to repair DSBs can result in genomic abnormalities, chromosomal rearrangements, and apoptosis. In mammalian cells, repair of DSBs proceeds by classical nonhomologous end joining (NHEJ), homologous recombination (HR), or by a third loosely described non-canonical repair pathway, termed alternative end joining (Alt-EJ). Amongst the most clearly defined characteristics of Alt-EJ is the presence microhomology, or small patches of complementary DNA sequence flanking a chromosome break, in the repair junction.

DNA Polymerase Theta (Pol θ) has been implicated as a primary mediator of this alternative repair pathway, but its cellular mechanism and role relative to canonical repair pathways is poorly understood. We show that Polymerase Theta mediated end-joining (TMEJ) accounts for most repair associated with microhomologies, and is made efficient by coupling a microhomology search to removal of nonhomologous tails and microhomology-primed synthesis across broken ends. In contrast to NHEJ, TMEJ efficiently repairs end structures expected after aborted homology-directed repair (5’ to

3’ resected ends) or replication fork collapse. It typically does not compete with canonical repair pathways, but in NHEJ-deficient cells is engaged more frequently and protects against translocation. Cell viability is also severely impaired upon combined deficiency in Pol θ and a factor that antagonizes end resection (Ku or 53BP1). TMEJ thus employs a flexible mechanism to help sustain cell viability and genome stability by rescuing chromosome break repair when resection is misregulated or NHEJ is compromised.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dale, for his guidance and support over the course of my doctoral work. Dale possess an incredible passion for science, an even more incredible wealth of knowledge, and has been integral in my development as a scientist. I would also like to thank my committee members Cyrus Vaziri, Jeff Sekelsky, Mike Emanuele, and Shawn Ahmed for both their time and interest in my work.

I am very grateful for all the Ramsden lab members that I have worked with over the past five years. Without the training, insights, and thoughtful discussion that they have provided, this work would not have been possible. In particular, I would like to thank Dr. John Pryor. His friendship and advice in navigating all aspects of the doctoral process has been invaluable.

I would not have arrived to this point without the support of my family and I am thankful for the opportunities that they have provided over the years. Finally, I would like to thank my fiancée Elizabeth Mutter-Rottmayer who has been an endless source of support, patience, and motivation.

Additional acknowledgments specific to each chapter

Chapter 1

This chapter was written entirely by D.W.W. Figure 1.2 was generated by M.J.Y, for publication in PLoS Genetics in 2014. D.W.W was a co-author on this manuscript.

Chapter 2

This chapter was modified from its original version appearing in Molecular Cell in 2016. I have divided the publication here for the purpose of logical flow and have modified figure layouts for this dissertation. Supplemental figures and 2 tables were included here. I have additionally included one piece of unpublished data (Figure 2.4). D.A.R., R.D.W, and D.W.W. designed the studies and analyzed the data. D.W.W. and M.J.Y. generated and characterized cell lines deficient in Pol θ and Ku70 and variants. D.W.W. performed all experiments presented in this chapter. D.W.W. and S.A.R. performed bioinformatic analyses. D.A.R. and D.W.W. wrote the manuscript with input from all authors. Studies presented in this chapter were supported by NCI grant CA097096 (D.A.R), AHA grant 14PRE20380258 (D.W.W), RP130297 and RP150102 from the Cancer Prevention and Research Institute of Texas and the Grady F. Saunders Research Professorship (R.D.W.), T32 CA09480 (MJY), DOD grant BC141727 and NIEHS grant R00ES02633 (S.A.R.), and the Burroughs Welcome Fund Career Award for Medical Scientists (G.P.G.). I thank Luis Blanco for the gift of *Polm*$^{-/-}$*Pol*$^{-/-}$ MEFs.

Chapter 3

This chapter was modified from its original version appearing in Molecular Cell in 2016. I have divided the publication here for the purpose of logical flow and have modified figure layouts for this dissertation. Supplemental figures and 3 tables were included here. I have additionally included one piece of unpublished data (Figure 3.4). D.A.R., G.P.G, R.D.W, and D.W.W. designed the studies and analyzed the data. D.W.W. and M.J.Y. generated and characterized cell lines deficient in Pol θ and Ku70

and variants. D.W.W. performed chromosomal end joining assays, resection assays, sequencing library preparation, translocation, and cell proliferation assays. W.F. performed immunofluorescence and colony formation assays. M.P.C. performed homologous recombination assays. D.W.W. and S.A.R. performed bioinformatic analyses. D.A.R. and D.W.W. wrote the manuscript with input from all authors. Studies presented in this chapter were supported by NCI grant CA097096 (D.A.R), AHA grant 14PRE20380258 (D.W.W), RP130297 and RP150102 from the Cancer Prevention and Research Institute of Texas and the Grady F. Saunders Research Professorship (R.D.W.), T32 CA09480 (MJY), DOD grant BC141727 and NIEHS grant R00ES02633 (S.A.R.), and the Burroughs Welcome Fund Career Award for Medical Scientists (G.P.G.). I thank Juan Carvajal-Garcia and George W. Small for methods development pertaining to this chapter.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

Alt-EJ, alternate end joining

bp, base pair

BRCA1/2, breast cancer 1/2

BRCT, BRCA1 Carboxy-Terminal

CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats

CtIP, CtBP Interacting Protein

ddPCR, digital droplet polymerase chain reaction

DNA-PKcs, DNA dependent Protein Kinase catalytic subunit

dRP, deoxy ribose phosphate

ds, double stranded

DSB, double strand break

EMSA, electrophoretic mobility shift assay

HR, homologous recombination

IR, ionizing radiation

kDa, kilo Dalton

MEF, mouse embryo fibroblast

MMBIR, microhomology-mediated break induced replication

MMEJ, microhomology-mediated end joining

MRN, Mre11-Rad50-Nbs1

NHEJ, nonhomologous end joining

nM, nanomolar

nt, nucleotide

PAGE, polyacrylamide gel electrophoresis

PARPi, Poly(ADP)ribose polymerase inhibitor

PBS, phosphate buffered saline

Pol, polymerase

PCR, polymerase chain reaction

qPCR, quantitative polymerase chain reaction

SNP, single nucleotide polymorphism

ss, single stranded

V(D)J, Variable, Diversity, Joining

ul, microliter

XLF, XRCC4-like factor/Cernunnos

XRCC4, X-ray cross-complementarity gene 4

γ, gamma

κ, kappa

λ, lambda

μ, mu

ν, nu

θ, theta

ζ, zeta

CHAPTER 1: INTRODUCTION

1.1 DNA double strand break repair

DNA double strand breaks (DSBs) are a rare but potentially lethal class of genomic damage. DSBs are generated by both exogenous agents (e.g. ionizing radiation (IR) and chemical mutagens) and during endogenous processes (e.g. meiotic recombination and V(D)J recombination). Efficient DSB repair is essential for maintaining genome integrity and cell viability, as unrepaired DSBs can induce apoptosis. On the organismal level, deficiencies in DSB repair display as phenotypes including accelerated aging, immunodeficiency, and cancer predisposition[1-5]. Induction of DSBs is the basis of many frontline cancer treatments, including IR therapy and chemotherapeutic drugs such as etoposide and bleomycin. These therapies rely on overwhelming the capacity of DSB repair processes or exploiting repair deficiencies in the tumor cell, ultimately resulting in cell death.

There are two major pathways for resolving DSBs: nonhomologous end joining (NHEJ) and homologous recombination (HR)[6] (Figure 1.1). In mammals, repair of DSBs proceeds predominantly by NHEJ[7]. NHEJ is a ligation pathway in which DSBs are resolved with little to no reliance on sequence homology at the breakpoint. This property of template independence allows NHEJ to be flexible with respect to damage at

DSB termini and to be robustly active throughout the cell cycle, at the expense of repair fidelity[8]. Frequently, repair by NHEJ results in small insertion/deletion events at the repair junction. Inaccurate repair of DSBs can induce mutations and chromosomal rearrangements that harbor the propensity to become oncogenic. In contrast to NHEJ, the other major repair pathway, HR, performs high fidelity repair of DSBs by using a sister chromatid or homologous chromosome as a template to direct repair. As such, HR is active only during S and G2 phases of the cell cycle when an undamaged template is available[9,10].

While NHEJ and HR are the predominant DSB repair pathways, a third, loosely described repair option termed both alternative end joining (Alt-EJ) and microhomology-mediated end joining (MMEJ) has been observed, primarily in systems that are deficient in classical NHEJ[11-15] (Figure 1.1). Alt-EJ/MMEJ is thought to be a "back-up" repair pathway, that while minor, may have significant effects on genome integrity that are not yet understood.

1.2 Nonhomologous end joining

The NHEJ core complex consists of the Ku heterodimer (Ku70:Ku80), DNA-PK_{CS}, XLF, and the DNA Ligase IV:XRCC4 complex[16]. NHEJ is initiated by the loading of Ku onto exposed DSB termini. The Ku heterodimer forms a ring structure that binds DNA with high affinity through a central channel and has the ability to translocate on DNA and peel DNA from chromatin[17-22]. In addition to direct roles in end processing, including 5'dRP lyase activity for the removal of nucleotide damage at the break termini, Ku serves as important hub for the recruitment of other NHEJ factors[23-25]. In the next step, DNA-PK_{CS} interacts with Ku at the site of the DSB to form the DNA-PK

holoenzyme. This large kinase forms a synaptic complex that bridges and aligns the broken ends[26]. Additionally, DNA-PK regulates access of processing factors such as polymerases and nucleases to the break ends, with autophosphorylation of DNA-PKcs mediating the remodeling of the synaptic complex to allow access[27,28]. Finally, the ligase complex consisting of DNA Ligase IV, with XRCC4 and XLF forming a stabilizing filament, ligates the aligned double strand break ends, thus restoring genome integrity[29-34]. DNA Ligase IV is a flexible translesion ligase that can tolerate mismatch distortions and some damage such as 8-oxo guanine nucleotides at the repair junction[35].

The termini of DSBs are rarely fully complementary, and frequently have associated nucleotide damage. Common damages include base mismatches, oxidized nucleotides, abasic sites (where the nitrogenous base is lost and only the sugar-phosphate backbone remains), and bulky adducts such as alkyl groups or covalently linked proteins[36-38]. A host of accessory processing factors associate with the NHEJ core complex to aide in the resolution of these damages. These factors include nucleases (e.g. Artemis), specialized factors for removing protein adducts (e.g. Tdp2), and X-family DNA polymerases (e.g. Pol λ and Pol μ)[39-42]. Of particular interest, these polymerases associate with the NHEJ core complex via BRCT domains and allow for alignment and gap filling synthesis of breaks that would otherwise require more extensive deletion to repair[43]. Pol λ and Pol μ are each individually active on cognate substrates and single loss of either enzyme is well tolerated by cells; yet double deficiency in both polymerases results in a sensitivity to ionizing radiation on-par with the loss of Ku[44]. Despite the incredible end-processing and ligation flexibility endowed to NHEJ, there arise times when damage is too severe, or cell cycle conditions dictate that

extensive nucleolytic processing of the DSB and resolution by a different pathway is preferred.

1.3 DNA end resection

DNA end resection is the 5' to 3' degradation of DSB termini, resulting in the generation of long 3' single-stranded DNA (ssDNA) tails. Resection is a tightly regulated process with a host of both pro- and anti-resection factors that control the fate of the DSB. Pro-resection factors include the MRN complex (MRE11, Rad50, and Nbs1), CtIP, and BRCA1[45-47]. A general model for resection involves the recognition of a break by the MRN complex, followed by nicking of the DNA by the endonuclease activity of MRE11 and 3'-5' degradation towards the break termini by the exonuclease activity of MRE11[48-52]. CtIP is recruited to the MRN complex by the Nbs1 component and is an essential accessory factor for Mre11 nuclease activity in this process[53-55]. BRCA1, though not essential for resection, is thought to increase the initial rate of resection through CtIP[56,57]. This initial processing results in relatively short 3' overhangs, on the order of 25-50 base pairs (bp). Subsequently, exonucleases and RecQ helicases such as ExoI, DNA2, and BLM are recruited to processively unwind and resect much larger tracts of DNA[58].

Proper regulation of resection is crucial for optimal repair pathway choice. Anti-resection factors include 53BP1, RIF1, PTIP, and Ku[59-64]. 53BP1 is recruited to DSBs through the reading of both methyl and ubiquitin marks on histones at sites of damage and serves as a scaffold for the recruitment of downstream anti-resection factors[65,66].

These factors suppress resection by physically blocking the accessibility of the DNA ends to pro-resection nucleases, thus promoting repair by NHEJ.

1.3.1 Initiation of homologous recombination

Repair by HR is essential for organismal development and is a major mediator of genetic diversity through meiotic recombination. The ssDNA tails exposed by end resection are an essential structural intermediate in the initiation of DSB repair by HR[67]. As the resection machinery exposes DNA tails, they are quickly coated by the trimeric ssDNA binding protein RPA and are protected from degradation by cellular nucleases and the formation of secondary DNA structures[68]. Subsequently, RPA is displaced from the ssDNA tails by the Rad51 recombinase. This exchange is promoted through BRCA2 and DSS1, which bind Rad51 monomers and destabilize RPAs affinity for ssDNA respectively, allowing Rad51 to properly bind and nucleate a protein:DNA filament along resected 3' ends[69-74]. The Rad51-coated nucleofilament then proceeds to invade the sister chromatid and search for homology between the resected end and the template chromosome[75]. Subsequently, a host of late-stage HR factors use the ssDNA tail as a primer to direct high-fidelity DNA synthesis off the intact template chromosome, and resolve the break via mechanisms that can result in either a crossover or non-crossover of genetic information between chromosomes.

1.4 Alternative end joining

The third DSB repair pathway, Alt-EJ/MMEJ, is poorly characterized relative to

both NHEJ and HR. This pathway has been defined primarily in the background of NHEJ deficiency and initially manifested as the observation that DSB repair junctions in the absence of NHEJ tended to join using small patches of complementary DNA ("microhomology") derived from sequence flanking the DSB[76,77]. Genetic analysis of this class of repair has provided evidence that Alt-EJ is dependent on the MRN complex and CtIP (and their orthologs MRX and Sae2, respectively, in *S. cerevisiae*), suggesting that like HR, Alt-EJ acts downstream of resection[78-80].

Very little is known about the cellular roles, context, and factors that are involved in the repair by of Alt-EJ. Alt-EJ appears to have some contribution to class switch recombination, where microhomology usage can occur during heavy chain rearrangement to generate different antibody isotypes, though a defined role in normal cellular repair remains elusive[81,82]. It has been suggested that Alt-EJ is dependent on the activity of Poly(ADP)ribose polymerase 1 (PARP1), as well as DNA Ligase I and DNA Ligase III[83-86]. Specifically, depletion of Ligase III leads to a decrease in microhomology-mediated chromosomal translocation junctions. These studies have tended to rely on the indirect output of reporter assays and bear further investigation, as there appear to be discrepancies between reporters both within and across species. Recently, studies have proposed a role for DNA Polymerase θ (Pol θ, Polq) in Alt-EJ/MMEJ.

1.5 DNA Polymerase θ

Pol θ is a 290 kDa protein encoded by the gene *POLQ.* Orthologs of this enzyme are found in most eukaryotes, but are not present in *S. cerevisiae*. Pol θ has a unique,

evolutionarily conserved domain structure consisting of a superfamily II helicase-like domain in the N-terminal third of the protein, a disordered central domain of unknown function, and an A-family polymerase domain in the C-terminal third[87]. The helicase-like domain does not appear to have any DNA unwinding activity, but does harbor an active ATPase motif[88]. Additionally, this domain harbors several putative Rad51 interaction motifs[89]. Recent structural studies of the helicase-like domain have proposed that the domain may serve as a ssDNA binding module, in which 3' overhang structures are channeled through a central region in the domain[90]. At the C-terminal end of the protein, the polymerase domain shares sequence identity with the other mammalian A-family polymerases, Pol ν and Pol γ, yet structurally diverges by the acquisition of 3 loop inserts[91]. Crystallization of this domain bound to DNA shows that these loop inserts make novel stabilizing contacts with DNA near the polymerase active site[92].

1.5.1 *In vitro* characterization of Polymerase θ

Pol θ has activity *in vitro* that is consistent with a relatively low fidelity DNA polymerase. The enzyme displays moderate processivity (up to 75 bp in a single round of extension), a single-base error rate (2.4×10^{-3}) similar to Y-family translesion polymerases such as Pol κ, and does in fact display some translesion synthesis activity[93]. For example, Pol θ can bypass thymine-glycol adducts and incorporate nucleotides across from abasic sites[94,95]. Further, Pol θ has the ability to extend following this base incorporation event, which could be biologically useful in the repair of damaged DNA. The polymerase also has a unique activity in that it can extend 3' ssDNA in both template-dependent and -independent manners, but only if the overhang

is far from duplexed DNA (at least 14 bp from duplex)[96]. This is opposite of X-family polymerases, which can perform template independent synthesis only when overhangs are short (< 4 nt). Finally, though the protein contains the evolutionary remnants of a 3'-5' exonuclease "proofreading" domain, it maintains no nucleolytic activity *in vitro*[92].

1.5.2 Polymerase θ and DSB repair

Several *in vivo* studies are strongly suggestive of a role for Pol θ in protecting genome stability, though there do appear to be subtle variations in Pol θ activities depending on species. The earliest studies of Pol θ as a mediator of genome stability came from a screen in which the *chaos1* allele, a mutation in mouse chromosome 16 that correlated with increased levels of micronuclei in blood samples (an indicator of chromosomal fragmentation), was determined to be a point mutation in Pol θ[97]. Further investigation of this phenotype by the generation of a full knockout mouse confirmed that loss of Pol θ was the causative agent of this instability. Of interest, mice made doubly deficient in Pol θ and ATM, a master regulator kinase of repair by HR, die during embryonic development, suggesting that Pol θ is involved in maintaining genome stability in a manner distinct from classical HR repair[98].

Pioneering studies in three model organisms (*D. melanogaste*r, *C. elegans*, and *M. musculus*) have reached the accord that Pol θ mediates a non-canonical DSB repair pathway, though with some evolutionary differences in cellular roles. The first direct reports of Pol θ involvement in a DSB repair pathway came from *D. melanogaster,* where end joining junctions following *p*-element excision only contained microhomologies in the presence of the Pol θ ortholog Mus308[99]. Further, these repair

junctions frequently contained templated insertion events that could be attributed to one or more microhomology dependent intra-molecular priming events[100]. *Drosophila* studies have also identified a role for Mus308 in the repair of inter-strand crosslinks (ICLs); a class of repair that typically involves a DSB intermediate[101].

Similarly, work in *C. elegans* has shown a crucial role for Pol θ in repair of complex replication blocks such as G-quadraplex DNA structures, CRISPR/Cas9 breaks, and damage induced by crosslinking agents[102-105]. Further, Pol θ appears to be a mediator of *C. elegans* germline mutagenesis and evolution through a mechanism that utilizes exactly one nucleotide of microhomology at DSB repair junctions[106].

Studies in mammalian systems have further confirmed a role for Pol θ in DSB repair, while highlighting evolutionary differences. Treatment of both WT and Pol θ-deficient murine bone marrow stromal cells with a variety of damaging agents showed that Pol θ confers a specific resistance to classical DSB inducing agents such as ionizing radiation and etoposide, but not to ICL inducing agents (Figure 1.2)[107]. These studies in mammalian systems have also addressed the broad role of Pol θ in genome stability through translocation formation. Translocations result from the aberrant joining of two DSBs on different chromosomes and are a common rearrangement in many cancers. Work in murine and human cells has suggested opposing roles for Pol θ in translocation formation (a suppressive role in mice compared with a promoting role in human), though the methods in each study varied substantially (spontaneous IgH:Myc translocation in mice compared with CRISPR/Cas9 induced breaks in human)[107,108].

1.6 Polymerase θ and cancer

Several types of cancer including stomach, lung, colorectal, and breast cancer have been associated with elevated Pol θ expression. Further, cancers overexpressing Pol θ are significantly associated with poor patient survival[109]. The relationship between Pol θ and breast cancer is currently the most explored association and in fact, expression levels of Pol θ and a cohort of its neighboring genes appear to be diagnostic for triple negative breast cancer[110]. Two single nucleotide polymorphisms (SNPs) in or near the *POLQ* gene have also been identified and associated with hereditary breast cancer in a case-control study of patients with either hereditary or sporadic breast cancers[111]. A SNP in the promoter region of the gene was most strongly associated, though the mechanistic implications of this SNP are unknown.

Recently, studies have suggested that Pol θ may be be a promising target in battling chemotherapeutic resistance. Co-inhibition of Pol θ and some HR-involved genes re-sensitized cisplatin-resistant lung cancer cells to the drug[112]. Similarly, inhibition of Pol θ greatly sensitized HR deficient tumor cell lines to a variety of cytotoxic agents including PARP inhibitors (PARPi). Most strikingly, relative tumor volume, clonogenic survival, and organismal survival were all positively impacted by depletion of Pol θ combined with PARPi therapy in a murine FANCD2-null xenograft model[89]. These results rationalize the exploration of Pol θ as a therapeutic candidate for the treatment of some cancers and highlight the need for a deeper mechanistic understanding of the enzyme and its contribution to cellular DSB repair.

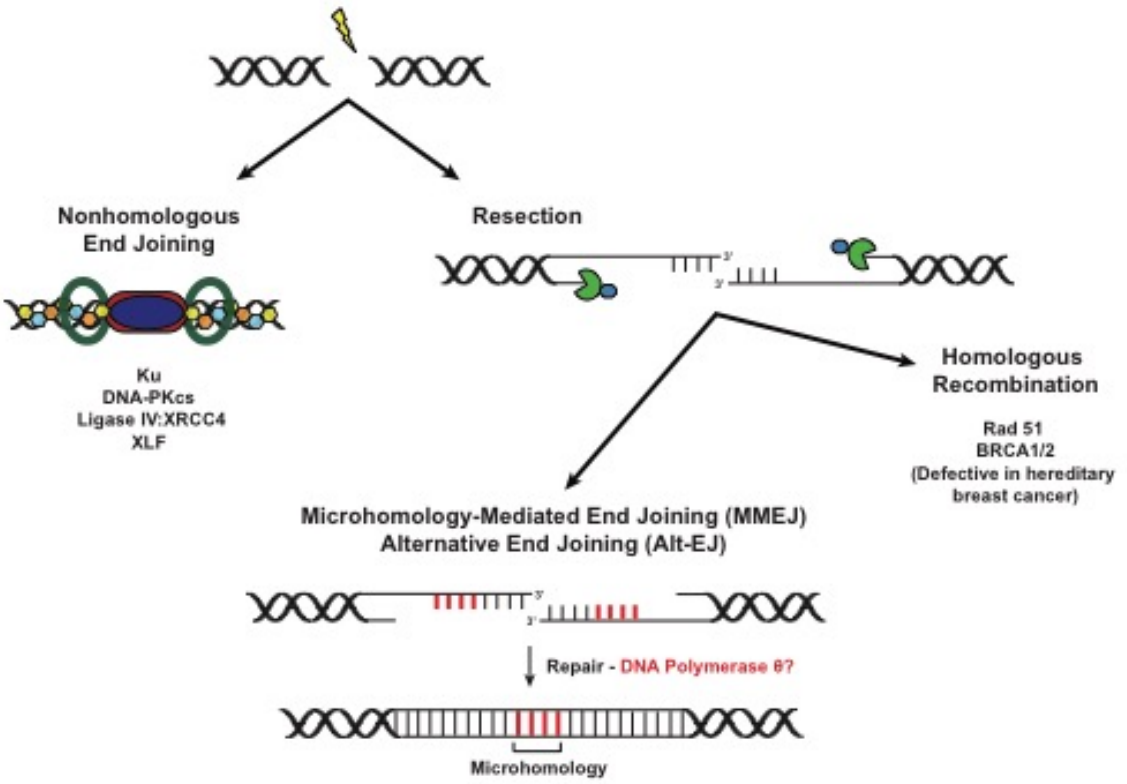

Figure 1.1: DNA double strand break repair pathways.

Cells repair DSBs through the nonhomologous end joining, homologous recombination, and alternative end joining/microhomology-mediated end joining pathways.

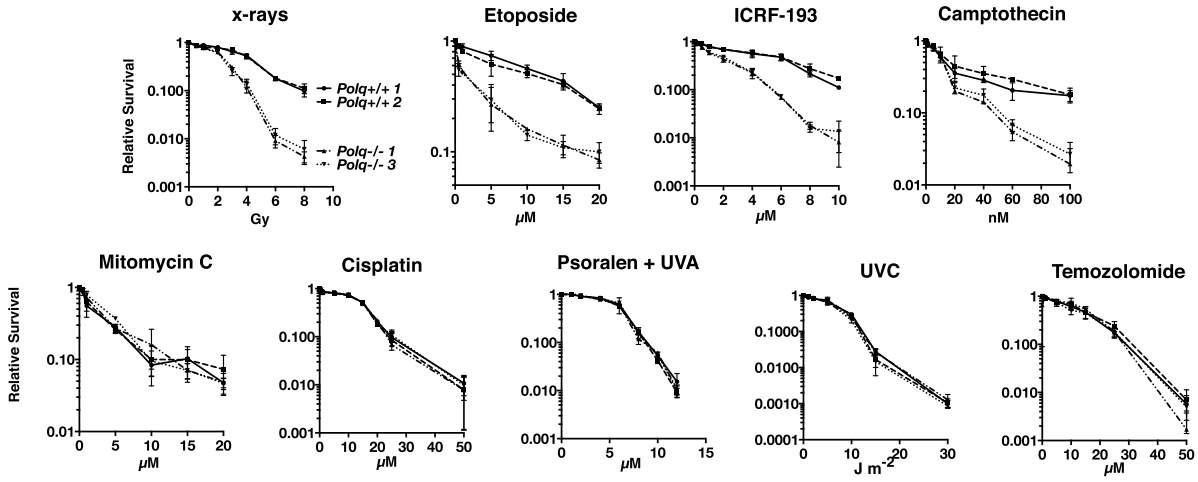

Figure 1.2: Hypersensitivity of *Polq*$^{-/-}$ bone marrow stromal cells to DNA strand-breaking agents.

BMSCs were exposed to x-rays or UVC at the indicated doses, and to etoposide, ICRF-193, camptothecin, olaparib, temozolomide, mitomycin c, cisplatin, and HMT psoralen+UVA at the indicated concentrations and plated in triplicate. Two isogenic bone marrow stromal cell lines were used of each genotype, Polq$^{+/+}$ or *Polq*$^{-/-}$. Colonies were stained and counted seven to ten days later.

CHAPTER 2: MECHANISM OF POLYMERASE θ-MEDIATED END JOINING

2.1 Introduction

DNA double-strand breaks (DSBs) can arise via a number of different exogenous exposures (e.g. ionizing radiation or chemotherapeutics), as well as through endogenous processes such as V(D)J recombination during immune system diversification and during normal cellular duplication cycles, when DNA replication stalls at naturally occurring structures or at sites of internally-generated DNA damage. Because DSBs are toxic and/or mutagenic if not repaired, organisms have multiple mechanisms for DSB repair[113,114]. The primary strategies are nonhomologous end joining (NHEJ), which processes and rejoins DSBs independent of sequence homology, and homologous recombination (HR) which employs an undamaged copy of the DNA for high fidelity repair[115]. In addition to these canonically studied repair processes, one or more "alternative" end joining pathways (Alt-EJ) also exist, which are independent of these factors[116,117].

Alt-EJ in eukaryotic cells has been linked to activity of DNA polymerase theta (Pol θ; encoded by *Polq*)[107]. This large, 290 kDa protein has an unusual bipartite structure with an N-terminal helicase-like domain and a C-terminal DNA polymerase domain[88]. This domain arrangement and the Pol θ protein sequence is highly conserved in vertebrates[87].

Several functions have been suggested for Pol θ, including bypass of template DNA lesions such as abasic sites and thymine glycols[95,118], a backup role in DNA base excision repair[119,120], and influencing the timing of DNA replication origin firing[121]. Loss of Pol θ homologs in *Drosophila* and *C. elegans* causes hypersensitivity to DNA interstrand crosslink (ICL)-forming agents[122,123] such as nitrogen mustards or cisplatin, while mammalian cells lacking Pol θ are uniquely sensitive to DSB inducing agents, such as ionizing radiation.

The mechanistic constraints on Pol θ activity in cellular repair are still unclear. We investigated these questions through use of extrachromosomal assays and a panel of mouse embryo fibroblast lines (MEF) cell lines deficient in *Polq* and *Ku70*. Our results provide insight into how Pol θ activity contributes to cellular DSB repair. We show Pol θ activity is facilitated by an ability to couple microhomology primed synthesis to a "microhomology search" and removal of non-homologous tails. We further identify an essential role for this enzyme in repair of DSBs in contexts where ends are inaccessible to canonical NHEJ, as is the case when resected ends exceed 45 nt in length, or at "Y" structures expected during replication fork collapse.

2.2 Methods

Cell lines

Experiments varying *Polq* status employed SV-40 Large T antigen transformed MEFs recovered from mice where a *Polq* null allele was generated by conventional gene targeting and made isogenic by backcrossing in mice. *Polq*$^{-/-}$ MEFs were then

complemented by lentiviral delivery of cDNA coding wild type human Polq. Deficiency or complementation was assessed by RT-qPCR for mRNA expression relative to a housekeeping control (GAPDH) (Figure 2.1A).

Variants of the above WT and *Polq*$^{-/-}$ lines deficient in the *Ku70* gene were made following transient expression of Cas9 and sgRNA targeting exon 3 of *Ku70*. Following Cas9 mutagenesis, cells were subcloned and screened for biallelic frameshift mutations. Complemented variants of these Ku70 deficient lines were made with lentivirus delivering mouse Ku70 cDNA. Ku70 deficiency or complementation was validated by western analysis (Figure 2.1B) and a functional assay for NHEJ activity (Figure 2.1C). All cell lines were cultured in DMEM with 10% FBS (Sigma), 2 mM L-Glutamine (Gibco), 1x penicillin-streptomycin, and 2 ug/ml puromycin, 4 ug/ml blasticidin, or 0.5 mg/ml G418 as necessary for transgene maintenance.

Substrate preparation

Substrates were prepared using a variation of the golden gate cloning method. A 596 base pair common core DNA segment was amplified by PCR and digested with BsaI-HF (NEB) to yield 4 nucleotide overhangs where head and tail overhang had distinct sequence. Caps with a common 15 bp double stranded region and overhangs specific for core head and tail ends were assembled by annealing of oligonucleotides (IDT), then ligating the caps to the core fragment, generate the linear fragments with head and tail end structures varied as noted in Table 2.1. Excess cap was removed using the QIAquick PCR purification kit (Qiagen) and substrate purity validated by polyacrylamide gel electrophoresis on a 4% gel.

Extrachromosomal end joining assay

DNA substrates (75 ng) of the noted substrate and pMAX-GFP (1 ug, Lonza) were introduced into $2x10^5$ cells by electroporation (Neon, Invitrogen) using a 1350 V, 30 ms pulse in a 10 µL chamber and incubated for 1 hour. Cells were washed, incubated with 12.5U of Benzonase (Sigma) to digest extracellular substrate, and cellular DNA purified (QIAamp DNA mini kit, Qiagen) to determine substrate joining efficiency and junction character by sequencing. Joining efficiency of extrachromosomal substrates introduced into cells was quantified by qPCR using either primers Sub Duplex Fwd/Rev or primers Sub Overhang Fwd/Rev, depending on the experiment. Total substrate recovery (both joined and unjoined) was determined as a validation of transfection efficiency across genotypes by qPCR using primers Sub Recovery Fwd/Rev. All PCRs were confirmed to be >98% efficient over the range relevant to these experiments by parallel analysis of a standard curve generated by serial dilution of model amplicons. Product structures were characterized by polyacrylamide gel electrophoresis on an 8% gel with SYBR green staining or sequencing.

Sequencing

Template DNA for each sequencing library was from independent electroporations performed across three days. $1x10^5$ input junction molecules were amplified using Phusion DNA polymerase (NEB) and variants of the qPCR primers noted above, with the addition of 6 nt barcode sequences appended to their 5' ends for 21 cycles. 15.5 ng of amplified DNA from each library was pooled in groups of ~10 libraries, phosphorylated with T4 PNK (NEB), treated with dATP and Klenow Exo – (NEB) to adenylate the 3' termini. Adaptors for sequencing (Illumina) were then

appended to the amplicons by ligation, and free adaptor removed by gel purification. Gel purified pooled libraries were subjected to a 10 cycle enrichment amplification with adaptor-specific primers and purified with Ampure XP beads (Agencourt, Beckman Coulter). Equal amounts of DNA from each pooled library set were combined and submitted for a 2x150 paired end sequencing run (MiSeq, Illumina), with a PhiX174 DNA spike.

After sequencing, reads of PhiX174 DNA were removed. Remaining reads were then trimmed for quality, paired ends were merged, and libraries were de-indexed using Genomics workbench v8.0 (CLC-Bio). Characterization of substrate repair junctions was performed by alignment of reads to a reference junction, export in SAM format, and deconstruction of the CIGAR string in Excel (Microsoft) to yield parameters including flanking deletion, length of ssDNA tails generated, and microhomology content at the junction.

Gel shift assay

Cy-5 labeled DNA substrates were generated by annealing of oligonucleotides to generate a 20 bp duplex core. Subsequent ligation of cap structures generated substrates with either 4 nt or 70 nt 3' ssDNA overhangs. Resulting substrates (5 nM) were incubated with 500 nM Streptavidin for 10 minutes at room temperature, then incubated with 5nM Ku for 5 minutes at 37°C in 1x EMSA buffer (20 mM Tris pH 8.0, 90 mM NaCl, 25 mM KCl, and 10% glycerol) at 37°C before addition of a 10-fold excess of unlabeled dsDNA and electrophoresis.

Strand displacement assay

Substrates bearing either 4 nt or 45 nt 3' overhangs were prepared as described above, with the variation that substrates now contained a mispaired BamHI restriction endonuclease site embedded 15 bp into the duplex DNA adjacent to the 3' overhangs. Substrate variations also contained either a 5' phosphate or 5' THF. The restriction site was mispaired such that 4 out of 6 positions of the BamHI recognition site were mispaired. Substrates were introduced into WT MEFs by electroporation as described above and harvested after 1 hr. Repair junctions were digesting with BamHI overnight at 37°C in a 20 ul reaction volume then used as template in a PCR amplification using primers Sub Duplex Fwd/Rev. Products were analyzed by PAGE electrophoresis on a 6% gel and qPCR.

2.3 Results

Distinct substrates for NHEJ and Pol θ-mediated end joining

We previously showed that Pol θ is required for a cellular end joining pathway – Polymerase θ mediated end joining, or TMEJ – that is independent of Ku, and uniquely able to join ends with extended 3'ssDNA tails[107]. To better address the relationship between TMEJ and classically defined NHEJ we sought to make cell lines deficient in one or the other pathway, or both. We consequently employed CRISPR/Cas9 to generate variants of existing isogenic wild type and *Polq*$^{-/-}$ MEF lines[107](Figure 2.1A) that do not express significant Ku70 protein (Figure 2.1B). These lines are deficient in NHEJ, and can be complemented by expression of introduced Ku70 cDNA (Figure 2.1C).

The progressive 5' to 3' resection of DSB termini is a necessary pre-requisite for repair of DSBs by homologous recombination (HR). To address how resection affects the deployment of different end joining pathways we introduced a panel of linear double stranded DNA substrates with "pre-resected" ends, where 3' ssDNA overhangs ranged from 4-70 nucleotides (nt) (Figure 2.2A, Figure 2.6A), into the cell lines described above. Head and tail overhang sequence were designed such that the terminal 4 nt were complementary, i.e. contain a microhomology. The impact of differing location and length of microhomologies is discussed in greater detail below.

Overall joining efficiency was reduced 5-10 fold for overhang lengths in excess of 10 nucleotides, and two populations of end joining products were readily evident; one where most of both overhangs had been removed, and another where most of both overhangs had been retained (Figure 2.2B-D). As overhang length is increased, joining is progressively more reliant on products where most of both overhangs are retained (Figure 2.6A, 2.6B). In all contexts joining after overhang loss was largely dependent on Ku70/NHEJ, while joining associated with retention of 3' overhangs was dependent on Pol θ (bottom panels, Figure 2.2B-D). Pol θ additionally was not required for either product class (overhangs retained and overhangs lost) using a substrate where the overhang polarity was reversed (5' overhang), but which was otherwise equivalent (Figure 2.6C). TMEJ is thus uniquely employed for repair of products of 5'>3' resection.

When the 3' overhang was only 45 nt, overall joining efficiency was reduced 3-fold in Ku70 deficient cells and not affected by Pol θ deficiency (Figure 2.2B, top panel). In contrast, the reciprocal pattern was observed when overhang length was increased to 70 nt (Figure 2.2C, top panel). Only trace levels of end joining were observed for both

substrate variants in cells deficient in both Pol θ and Ku70 (Figure 2.2B, 2.2C, top panels). The two pathways thus together account for most of the cellular capacity to repair pre-resected end structures, but TMEJ assumes prominence as the extent of 5'>3' resection exceeds 45 nt (see also Figure 2.6B).

We reasoned that this increased reliance on Pol θ-dependent repair as the ssDNA tails are extended reflects a reduced ability of the Ku heterodimer to load on substrates with long ssDNA tails (Figure 2.6C). To further investigate this possibility, we generated a variant of the shorter 45 nt 3' overhang with a "Y" or forked end structure, by including a 5' streptavidin-blocked flap (Figure 2.2D). Such a substrate is expected to completely block loading of Ku, and is analogous to products of replication fork collapse ("one ended breaks"). Accordingly, we observed no significant impact of Ku-dependent NHEJ when using this substrate. Repair was instead Pol θ-dependent (Figure 2.2D, top panel), consisted almost entirely of products that retained the majority of both 3' overhangs (Figure 2.2D, bottom panel), and thus can be readily defined as TMEJ. We conclude Pol θ-dependent repair is most important when considering end structures where loading of Ku is impaired. This includes those end structures expected after aborted HR – ends with extended 3' ssDNA tails - or after replication fork collapse.

Mechanism of Pol θ-mediated end joining

Products that retain significant amounts of overhang are mostly dependent on Pol θ; conversely, Pol θ deficiency had little impact on either the efficiency or structure of those products generated after overhang loss (Figure 2.2). Relocation of qPCR primers to require that end joining products retain at least 10 nt of each overhang is thus sufficient to restrict analysis to Pol θ-mediated events[107](Figure 2.3A).

Pol θ is a multi-domain protein, with an N-terminal helicase-like domain and a C-terminal polymerase domain (Figure 2.7A). To investigate the contributions of each activity we compared the ability to complement *Polq*$^{-/-}$ cells for TMEJ capacity by expression of the wild type *Polq* cDNA, a variant cDNA with a mutation inactivating ATPase activity of the helicase domain (K121M), or a variant cDNA with mutations that inactivate the polymerase domain (D2330A/Y2331A; Motif A)[107]. Levels of TMEJ in cells expressing the D2330A/Y2331A construct were indistinguishable from cells with an empty vector control (Figure 2.3A); polymerase activity is thus essential for this pathway. In contrast, loss of ATPase activity in the helicase-like domain had no significant impact on joining efficiency, but did affect product spectra (discussed below). We also assessed a possible role for Pol μ and Pol λ, which have previously been associated with an ability to direct intermolecular DNA synthesis, i.e. can prime synthesis from a 3' tail from one DSB end, and use a second DSB end as template[44]. Both of these polymerases are dispensable for cellular end joining using pre-resected substrates (Figure 2.3A), with levels of TMEJ instead increased in cells deficient in both Pol μ and Pol λ. This may reflect the important role of these polymerases in NHEJ, and increased TMEJ in NHEJ-deficient contexts.

The substrates described above were designed such that the terminal 4 nt were complementary, to allow for alignment-directed synthesis. Surprisingly, products consistent with synthesis from these terminal alignments accounted for only 56% of TMEJ when using the 45 nt 3' overhang substrate, and 20% of TMEJ when using the 70 nt 3' overhang substrate (Figure 2.3B; Table 2.2). Remaining products favored alignments with short complementary sequences found internally ("embedded"

alignments). For example, repair of the 70 nt overhang substrate frequently employed all four of the alignments that are greater than 3 bp and can also be found within 25 nt of the 3' ends. Taken together, these 4 products account for 2/3rds of all products recovered (Figure 2.3B), with the three embedded alignments used almost twice as often as the terminal alignment. Moreover, there were only mild effects on the efficiency of repair when using a substrate where both the terminal and most favored embedded alignments were disrupted (Figure 2.7B). We additionally did not see significant recovery of products where an embedded alignment directed synthesis from a mispaired tail (Figure 2.3B). We can therefore infer synthesis by Pol θ using embedded alignments is efficiently coupled to a prior step where the unpaired tail is removed.

Notably, the ability to use embedded alignments was modestly reduced in cells expressing the Pol θ variant with an ATPase inactivating mutation (K121M; Figure 2.3C). This suggests TMEJ is facilitated by an active search for complementary sequence alignments (i.e. microhomology in finished product) that involves ATPase activity in the N-terminal helicase-like domain. Characterization of products also clarifies the constraints on this putative microhomology search mechanism. Considering first the amount of complementary sequence, only those products with alignments > 2bp were significantly enriched (Figure 2.3D). However, the length of nonhomologous tails impacts the ability to use such alignments. For example, a product generated from a 6 bp alignment, but which would generate long nonhomologous tails (52 nt and 22 nt), is observed 20-fold less frequently than a 4 bp alignment with short tails (9 nt and 2 nt, Figure 2.3B). Alignments resulting in tails with one long and one short tail (e.g. 51 nt and 4nt, Figure 2.3B) are also rarely used, suggesting a need to recognize both 3'

termini. In sum, 91% of products recovered using the 70 nucleotide overhang substrate (140 nt of ssDNA) have deletion less than 50 bp (Table 2.2).

Moreover, we show here that the synthesis activity of TMEJ does not terminate at a downstream 5' phosphate and instead involves a strand displacement intermediate. Repair of substrate containing a BamHI site on the template strand that is mispaired and buried 15 bp into the substrate duplex is not sensitive to restriction enzyme digest when overhang length is 4 nt and thus repaired by NHEJ (Figure 2.4A, 2.4B). Joining of this substrate is reduced when the 5' phosphate is replaced with a 5' abasic site mimetic (e.g. THF). In contrast, when the overhang length is extended to 45 nt, the substrate is repaired by TMEJ and junctions are rendered sensitive to pre-amplification digestion by BamHI (Figure 2.4B). Further, joining of this substrate not influenced by the presence of a 5' abasic site in lieu of a 5' phosphate. This result confirms that the synthesis step of TMEJ preferentially displaces the mispaired strand and uses the intact BamHI-containing strand as template in a "long patch" DSB repair process.

2.4 Discussion

Our results indicate that in cells, Pol θ has a central role in an end joining pathway – TMEJ – that is uniquely able to repair substrates where classically defined NHEJ is ineffective (Figure 2.2). The effects of Pol θ deficiency on both which substrates can be joined (Figure 2.2) and product spectra (Figure 2.3, 2.4) indicates that in mammals, Pol θ/TMEJ is required for most of what had previously been defined as MMEJ or Alt-NHEJ. Characterization of these products further argues TMEJ is an organized pathway, as

synthesis by Pol θ is coupled to two prior steps, including a search for microhomology and removal of nonhomologous tails (Figure 2.3).

Identification of biologically significant substrates and mechanism

We show there are marked differences between cellular TMEJ and the in vitro activities of the truncated Pol θ polymerase domain. TMEJ is uniquely required as ssDNA extensions exceed 45 nt (Figure 2.2), as well as forked ends (i.e. when a paired 5' terminus is unavailable); by comparison, *in vitro* studies[124] indicate activity was limited to extensions 15 nt or shorter. We suggest full length Pol θ may be intrinsically more processive[93], Pol θ may employ a processivity clamp (PCNA) during cellular TMEJ, or there may be polymerase switching after synthesis initiation. The unique activity on forked structures is consistent with studies identifying an important role for Pol θ in replication fork re-start after fork collapse ("one-ended" chromosome breaks)[89,103]. Indeed, this result predicts TMEJ will be critical in joining any pair of ends where either of the 5' termini is inaccessible to Ku or "unligatable", i.e. has irreversible bulky adducts or is associated with extensive base damage. In accord with this prediction, we show that TMEJ is not impeded by a 5' abasic site – a lesion that ablates joining by NHEJ (Figure 2.4).

Cellular TMEJ is also rarely associated with synthesis from terminal mispairs (Figure 2.3), as can be observed in in vitro studies of truncated Pol θ activity[124]. This limitation could dramatically restrict the fraction of ends that could be joined (<1/64), given the requirement for annealing at significant microhomologies (3-6 bp) for synthesis. Our results indicate cellular TMEJ instead efficiently samples all of the useable alignments within the 3' terminal 25 nt of a pair of ends, which is sufficient to enable efficient joining

of over 99% of ends of random sequence. The spectrum of cellular TMEJ products also emphasizes another advantage of a homology-search mechanism, its ability to mitigate deletion. 91% of TMEJ-dependent products of an extrachromosomal substrate with 140 nucleotides of ssDNA have less than 50 bp (2x25) of deletion.

Cellular TMEJ is thus made flexible, and deletion constrained, through an ability to tightly couple together three steps (Figure 2.5): 1) a microhomology search that usually generates nonhomologous tails, 2) excision of these tails, and 3) strand-displacement synthesis from the newly annealed 3' terminus. The search for microhomology might be dependent on the Pol θ N-terminal helicase-like domain, as suggested by a structural model for an oligomer of this domain that juxtaposes a pair of 3' ssDNA termini[90]. In accord with this model, the ability to use embedded alignments is specifically impaired in cells expressing a Pol θ mutation that inactivates this domain's ATPase activity. Effects of ATPase activity on microhomology search could plausibly be explained if it drove adjustment of the relative position of the two aligned termini. With regard to the next step, removal of nonhomologous 3' tails, Pol θ has no intrinsic nuclease activity[87]. However, such an activity is present as a contaminant of Pol θ partially purified from HeLa cells[125], and results described here suggest this activity is tightly coupled to Pol θ activity. Mre11 and XPF:ERCC1 are both attractive candidates for this nuclease activity, as they have appropriate substrate specificities and have already been implicated in Alt-NHEJ/MMEJ[77]. The third step, synthesis from 3' termini with 3-6 bp annealed, is facilitated by the tight grasp Pol θ maintains on the primer termini[92]. Altogether, Pol θ is uniquely suited to direct this pathway.

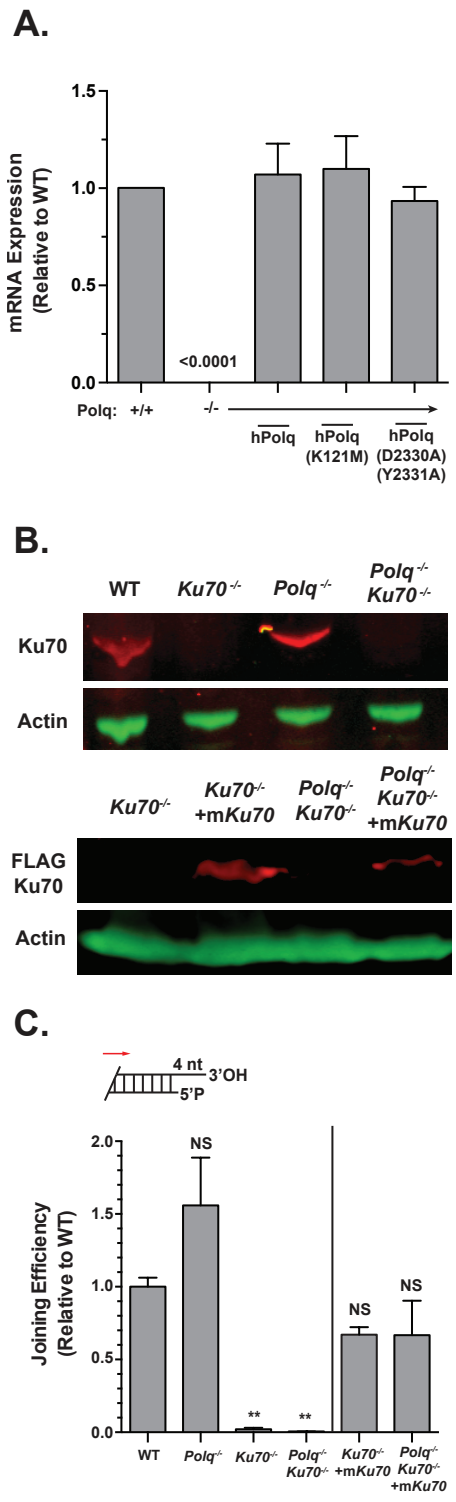

Figure 2.1: Validation of NHEJ deficient cells and complementation

Figure 2.1: Validation of NHEJ deficient cells and complementation

(A) Levels of endogenous Polq or hPolq mRNA were assessed by a qPCR of reverse transcribed RNA recovered from the noted cell lines using primers that have equal sequence identity to both human and mouse Polq cDNA. Data are the mean +/- SEM, n=3, relative to the amount observed in WT (*Polq*$^{+/+}$) MEFs. **(B)** Western blot of noted cell types assessing expression of endogenous Ku70 or introduced FLAG-tagged Ku70 proteins. **(C)** A c-NHEJ substrate with head and tail 4 nt 3' overhangs was introduced into noted cell types, and joining efficiency was assessed as described in Figure 2. Data are the mean +/- SEM, n=3. Statistical significance assessed by one-way ANOVA with Bonferoni correction.

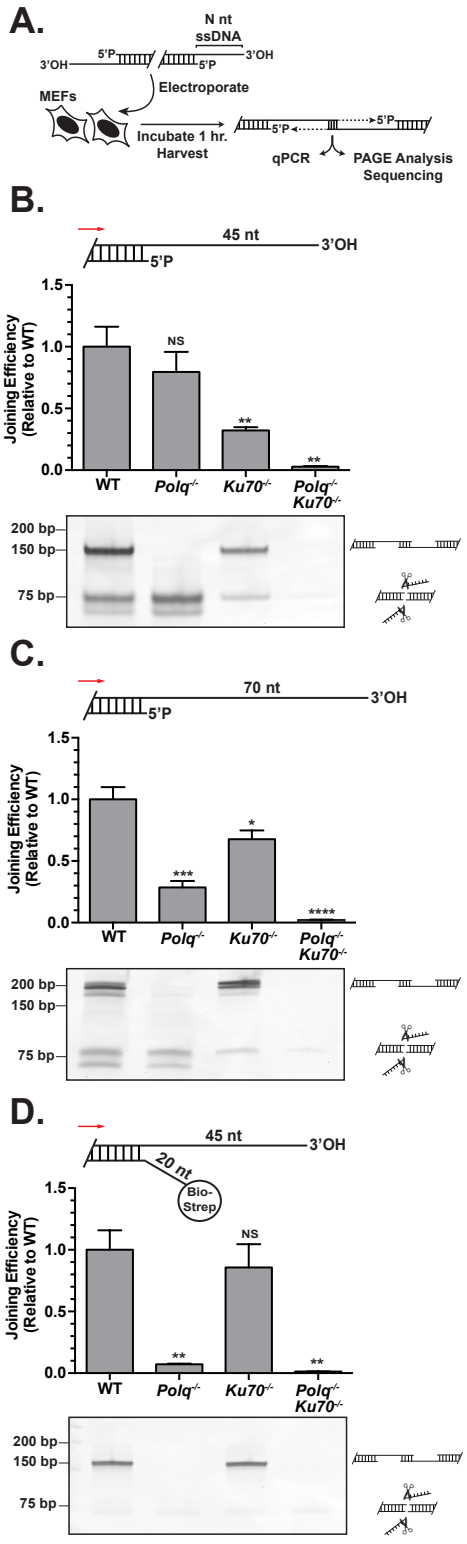

Figure 2.2: Effect of end joining deficiencies on repair of pre-resected ends

Figure 2.2: Effect of end joining deficiencies on repair of pre-resected ends

(A) Linear substrates with varied end structures were introduced into noted cell types and end joining products characterized by quantitative polymerase chain reaction (top panels) or electrophoresis of amplified products (bottom panels); Data shown below are the mean +/- SEM, n=3. Statistical significance was assessed by one-way ANOVA with Bonferoni correction of *p* values to account for multiple comparisons. NS not significant, **p*<0.05, ***p*<0.01, ****p*<0.001, *****p*<0.0001. **(B)** A substrate with symmetrical 45 nt 3' single stranded DNA overhangs was introduced into the noted cell types, and the mean efficiency of end joining determined by qPCR and expressed as a fraction of that observed in WT cells. Amplified products generated in a parallel fixed-cycle number PCR were also characterized by electrophoresis; products of size consistent with joining after overhang clipping vs. overhang retention are noted. **(C)** A substrate with 70 nt 3' overhangs was introduced into the noted cell lines and joining characterized as described above. **(D)** A substrate as in panel B, except with a 20 nt non homologous tail ending in a 5' terminal biotin-streptavidin group, was introduced into the noted cell lines and joining characterized as described above.

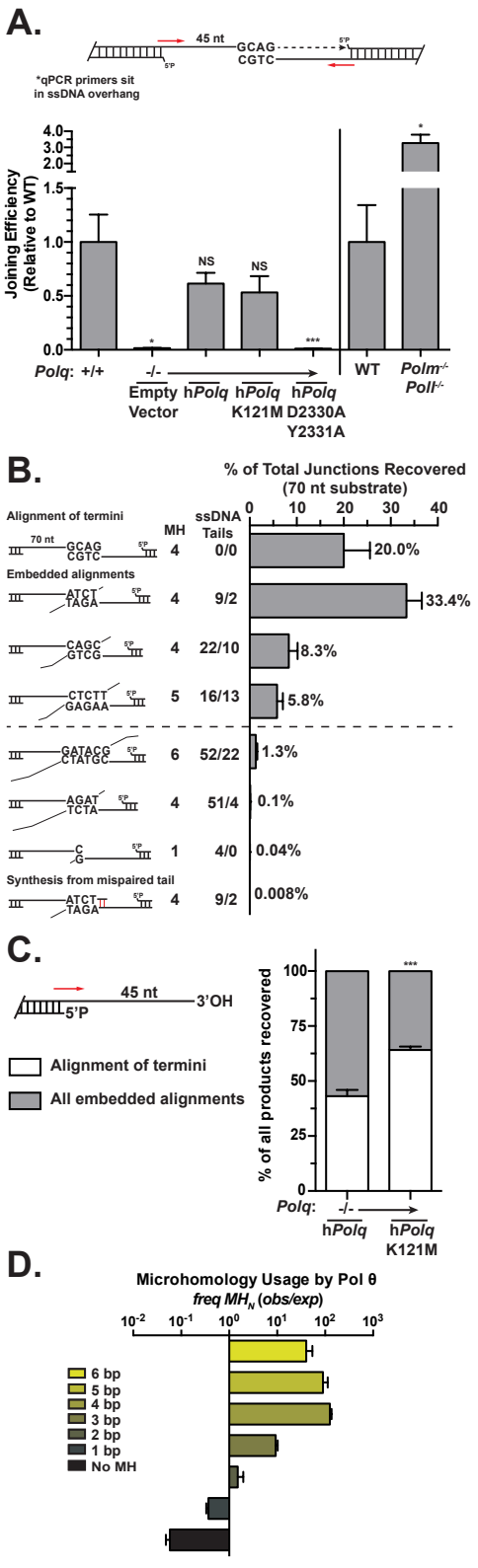

Figure 2.3: Mechanism of Pol θ mediated end joining

Figure 2.3:Mechanism of Pol θ mediated end joining

(A) The 45 nt 3' overhang substrate described in Fig. 2B was introduced into cells that were WT, *Polq*$^{-/-}$, and *Polq*$^{-/-}$ engineered to express wild type human *Polq* or variants deficient in helicase-like domain ATPase activity (K121M), or polymerase activity (D2330A+Y2331A), and joining efficiency assessed by qPCR using primers that require retention of at least 10 nt of each 3' overhang sequence in junctions. Joining efficiency was also characterized in *Poll*$^{-/-}$ *Polm*$^{-/-}$ MEFs, relative to matched WT control MEFs. Data shown are the mean +/- SEM, n=3. Statistical significance was assessed by one-way ANOVA with Bonferoni correction. NS =not significant, *p<0.05, **p<0.01, ***p<0.001. **(B)** Observed frequency of different joining products of the 70 nt overhang substrate (Fig. 2C) recovered from wild type cells, including the 4 most common products (above dashed line) and 4 other representative examples (see also Table S1). The first 3 columns summarize the structures of the inferred intermediates. **(C)** For experiments described in Fig. 3A (using the 45 nt overhang) we report the fraction of products directed by the terminal 4 bp microhomology (shown in panel B), as determined by sequencing. Data shown are the mean +/- SEM, n=3. Statistical significance was assessed by two-way ANOVA (p values as above). **(D)** The extent each sized microhomology is enriched (bars right of Y axis) or depleted (bars left of Y axis) in recovered junctions, relative to their representation in the set of all possible deletion products. See also Figure S3, Table S1.

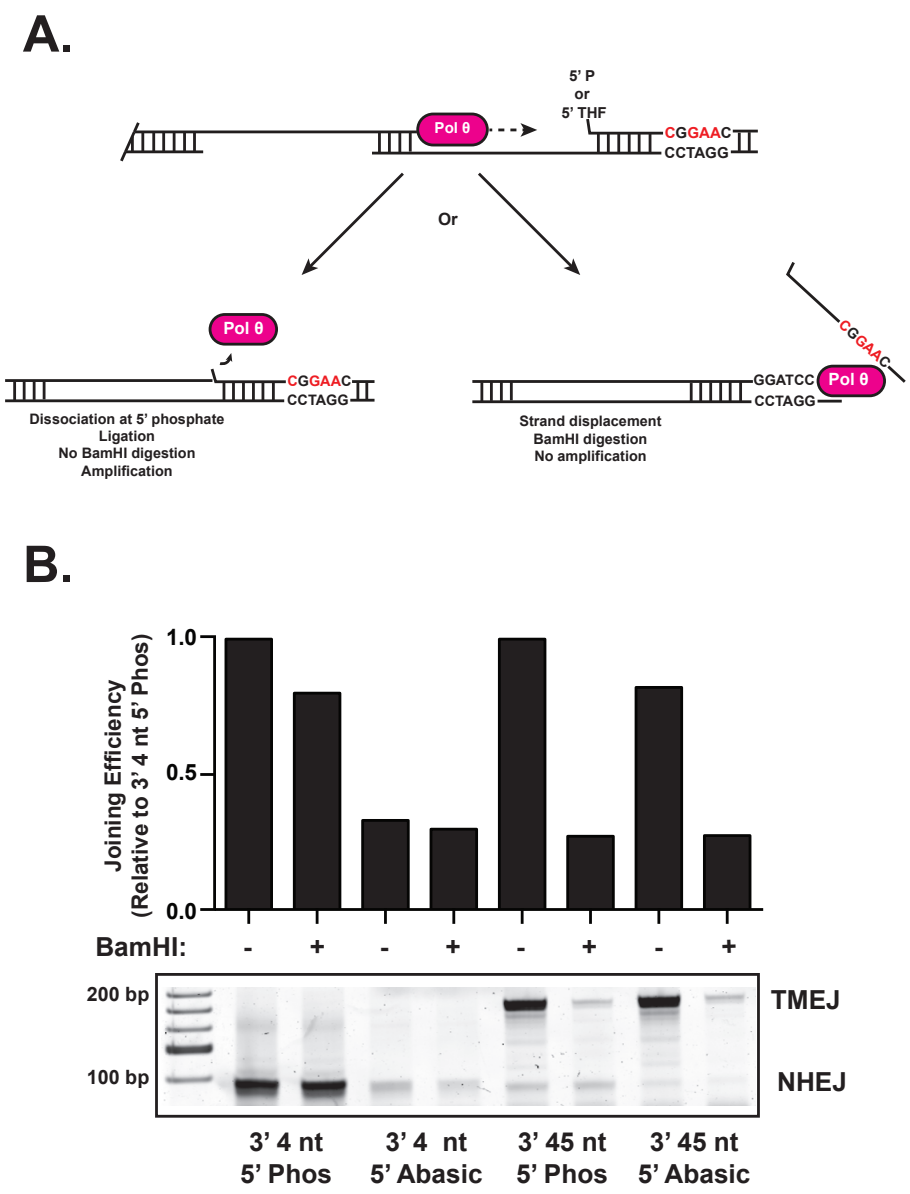

Figure 2.4: Pol θ is a strand displacing polymerase

(A) Schematic showing the possible Pol θ repair fates and the impact of each fate on digestion with BamHI and subsequent PCR amplification. **(B)** qPCR quantification (top panel) and 6% PAGE gel analysis (bottom panel) of noted substrates, either treated as a mock sample or BamHI digested, followed by PCR amplification.

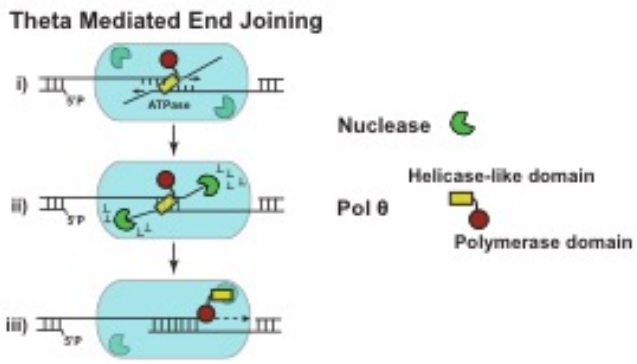

Figure 2.5: Proposed model of Pol θ-mediated end joining

Proposed mechanism of cellular Pol θ activity including a i) search for microhomology, ii) removal of non-homologous tails, and iii) synthesis.

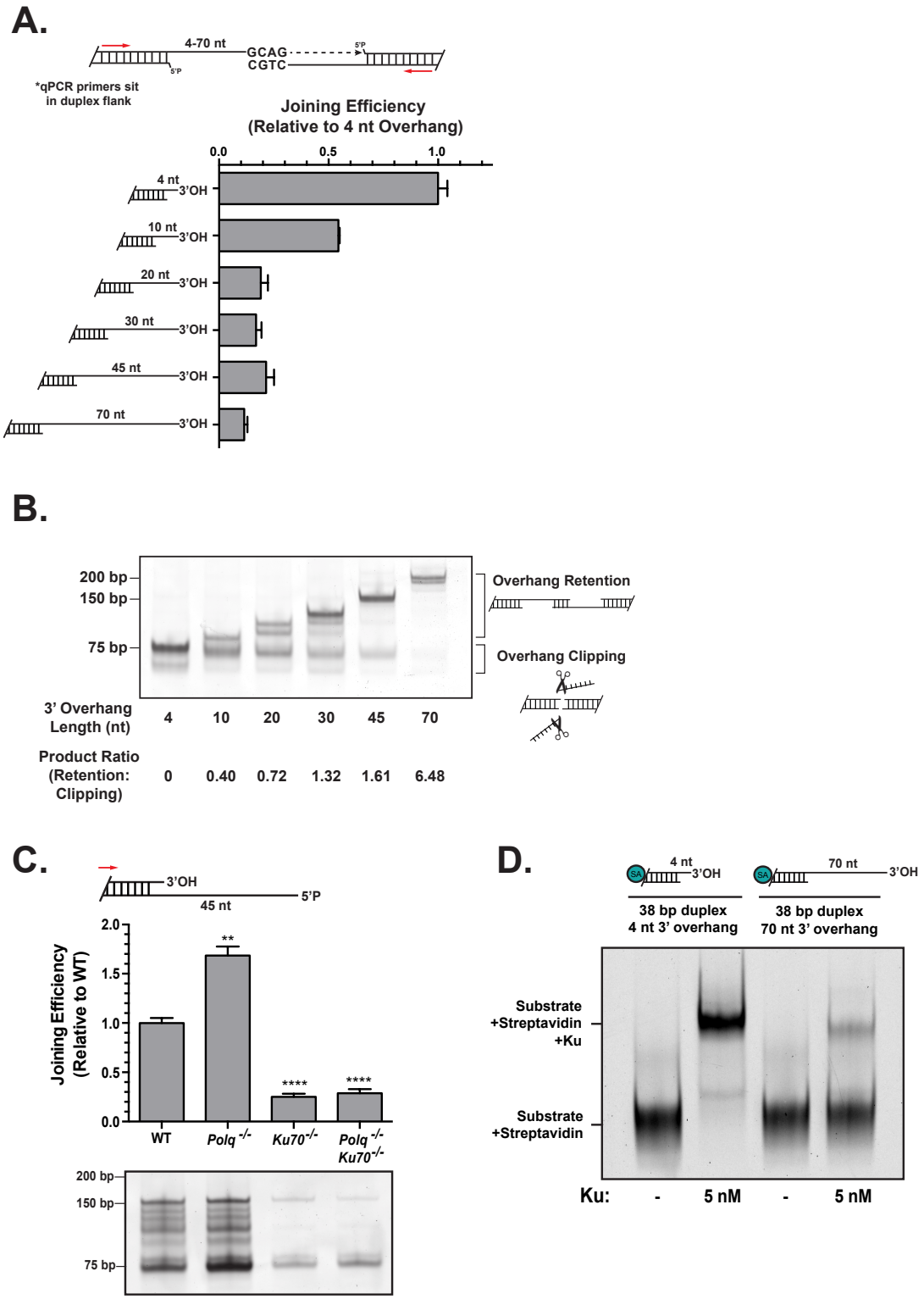

Figure 2.6: Effect of varied overhang length and polarity on end joining

Figure 2.6: Effect of varied overhang length and polarity on end joining

(A) Substrates with 3' single stranded DNA overhangs ranging from 4-70 nt were introduced into WT MEFs. Joining efficiency was quantified as described in Figure 2. Data are the mean +/- SEM, n=3. **(B)** Joined products from panel A were characterized by amplification of end joining products and polyacrylamide gel electrophoresis. Products of size consistent with joining after overhang clipping vs. joining with overhang retained are noted, and their relative abundance (determined after correction for size) noted below. **(C)** A substrate with head and tail 45 nt 5' overhangs was introduced into the noted cell types and the efficiency (top panel) and character (bottom panel) of end joining were assessed as described in Figure 2. Data are the mean +/- SEM, n=3. Statistical significance assessed by one-way ANOVA with Bonferoni correction. **(D)** Electrophoretic mobility shift assay showing loading of Ku onto substrates containing a 38 base pair duplex DNA blocked on one end with a biotin-Streptavidin moiety, and either a 4 nt or 70 nt 3' ssDNA overhang.

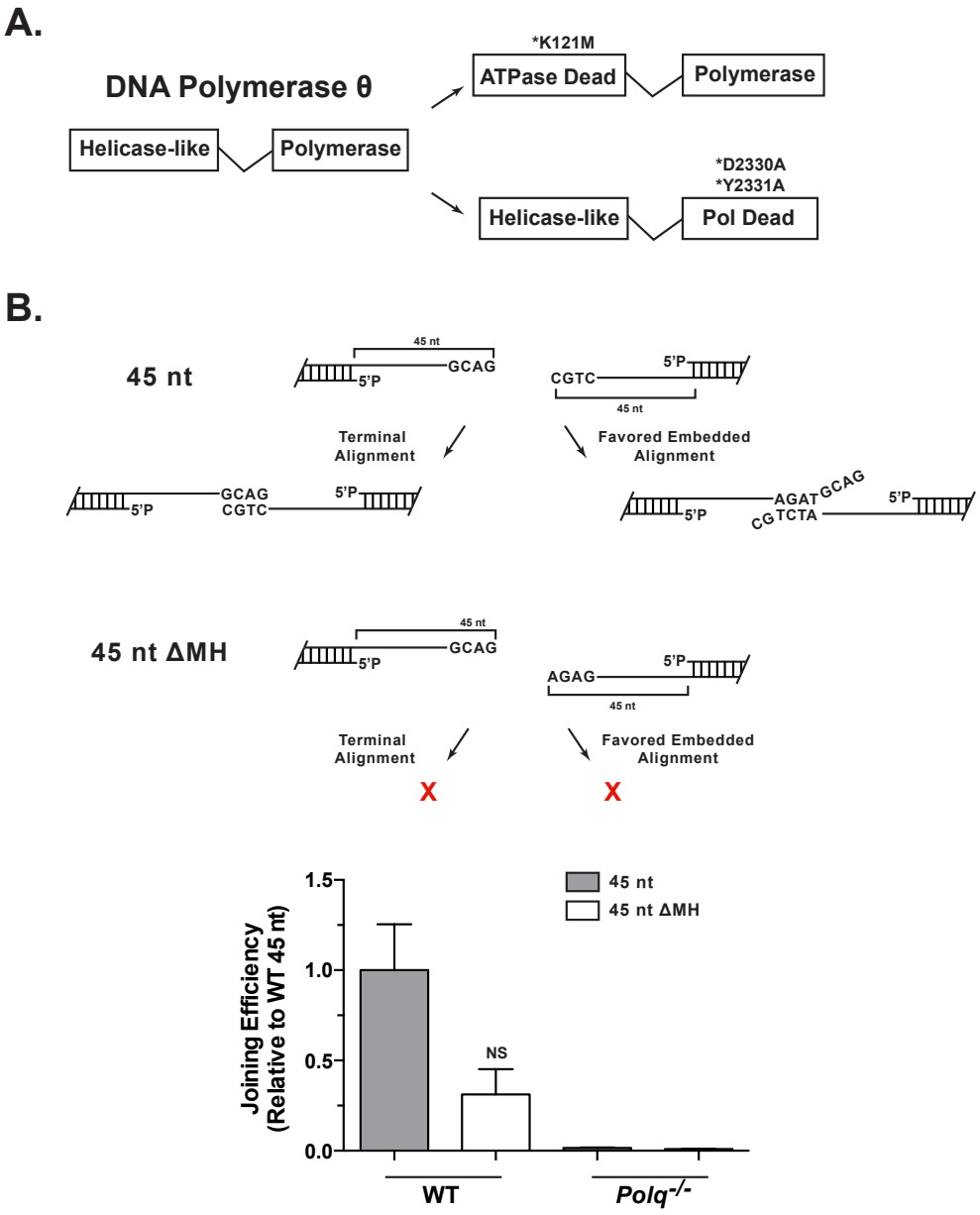

Figure 2.7: Separation of function in Pol θ and role of terminal microhomology

(A) Domain structure of Pol θ. Experiments performed in Figure 3 use *Polq*$^{-/-}$ MEFs expressing WT human Polq cDNAs, as well as Polq cDNAs with an ATPase defective mutation (K121M) or mutations that ablate polymerase activity (D2330A, Y2331A). **(B)** The 45 nt 3' overhang substrate described in Fig. 2B was transfected into WT and *Polq*$^{-/-}$ MEFs. A variant of this substrate with no ability to align either at the extreme 3' termini, or the second most favored alignment (shown in the cartoon) was transfected in parallel; joining efficiency was assessed using overhang specific primers as described in Figure 3. Data are the mean +/- SEM, n=3. Statistical significance assessed one-way ANOVA with Bonferoni correction.

Extrachromosomal Substrate: Left Caps	
4 nt	5’P/AGTCTGAGATGGGTGCCACGACG
10 nt	5’P/AGTCTGAGATGGGTGTGAGAGTGATCTGC
20 nt	5’P/AGTCTGAGATGGGTGTGAGAGTGAAGATCCTCGATCTGC
30 nt	5’P/AGTCTGAGATGGGTGTGAGAGTGAAGATCCTCACCTTCGGACGATCTGC
45 nt	5’P/AGTCTGAGATGGGTGTGAGAGTGAAGATCCTCACCTTCGGAGTACTCCTTCTTTTGAGATCTGC
70 nt	5’P/AGTCTGAGATGGGTGTGAGAGTGAAGATCCTCACCTTCGGAGTACTCCTTCTTTTGACCATTGATACGATACTTCTCAGCCGAGATCTGC
45 nt ΔMH	5’P/AGTCTGAGATGGGTGTGAGAGTGAAGATCCTCACCTTCGGAGTACTCCTTCTTTTGACTAGAGA
45 nt 5’ Overhang	5’P/GCAGATCTCAAAAGAAGGAGTACTCCGAAGGTGAGGATCTTCACTCTCACACCCATCTCA
Annealing Strand	5’P/CTCACACCCATCTCA
5’ Biotin Flap	5’Biotin/TTTTTTTTTTTTTTTTTTTTCTCACACCCATCTCA
Annealing Strand for 5’ Overhang	5’P/AGTCTGAGATGGGTGTGAG

Extrachromosomal Substrate: Right Caps	
4 nt	5’P/TGACTATACAGCTAAGCCCACGACG
10 nt	5’P/TGACTATACAGCTAAGCGATGCTCGATGCAG
20 nt	5’P/TGACTATACAGCTAAGCGATGCTCTCACCGAGCGGATGCAG
30 nt	5’P/TGACTATACAGCTAAGCGATGCTCTCACCGAGCGTATCTGCTGTGATGCAG
45 nt	5’P/TGACTATACAGCTAAGCGATGCTCTCACCGAGCGTATCTGCTGTGTTGTGGATGAATTAGATGCAG
70 nt	5’P/TGACTATACAGCTAAGCGATGCTCTCACCGAGCGTATCTGCTGGGTTGTGGATGAATTACATATGCTGGGAGAACCAAGATTGGATGCAG
45 nt ΔMH	5’P/TGACTATACAGCTAAGCGATGCTCTCACCGAGCGTATCTGCTGTGTTGTGGATGAATTAGATGCAG
45 nt 5’ Overhang	5’P/CTGCATCTAATTCATCCACAACACAGCAGATACGCTCGGTGAGAGCATCGCTTAGCTGTATA
Annealing Strand	5’P/CATCGCTTAGCTGTATA
5’ Biotin Flap	5’Biotin/TTTTTTTTTTTTTTTTTTTTCATCGCTTAGCTGTATA
Annealing Strand for 5’ Overhang	5’P/TGACTATACAGCTAAGCGATG

Primers/Probes	
Sub Duplex Fwd	5’CTTACGTTTGATTTCCCTGACTATACAG
Sub Duplex Rev	5’GCAGGGTAGCCAGTCTGAGATG
Sub Overhang Fwd	5’TAAGCGATGCTCTCACCGAG
Sub Overhang Rev	5’GATGGGTGTGAGAGTGAAGATC
Sub Recovery Fwd	5’GGCACTCTCCAAGGCAAAGA
Sub Recovery Rev	5’ACATGTCTAGCCTATTCCCGGCTT

sgRNA Sequences	
Ku70 Exon 3	5’TCTTACTGGTGTACACACTC

Gel Shift Assay Oligos (Fig. S2)	
Common Duplex I	5’P/GTCACCCAAATCAAACGTAAG/TEGBiotin
Common Duplex II	5’Cy5/CTTACGTTTGATTTCCC
4 nt Overhang I	5’P/GTGGGCTTAGCTGTATA

4 nt Overhang II	5'P/TGACTATACAGCTAAGCCCACGACG
70 nt Overhang I	5'P/CATCGCTTAGCTGTATA
70 nt Overhang II	5'P/TGACTATACAGCTAAGCGATGCTCTCACCGAGCGTATCTGCTGGGTTGTGGATGAATTA CATATGCTGGGAGAACCAAGATTGGATGCAG

Antibodies	
Ku70	Abcam ab3114
FLAG Tag	Sigma F1804
Actin	Sigma A2066

Plasmids	
FLAG-Cas9	A gift of Xingxu Huang (Addgene plasmid # 44758)
sgRNA Plasmid	A gift of Xingxu Huang (Addgene plasmid # 51133)
pMAX-GFP	Lonza

Table 2.1: Reagents used during experiments.

45 nt Overhang			
Left Flap	**Right Flap**	**MH**	**Frequency**
0	0	4	0.560
4	2	4	0.301
11	3	3	0.013
3	12	3	0.012
2	4	1	0.011
13	34	4	0.005

70 nt Overhang			
Left Flap	**Right Flap**	**MH**	**Frequency**
9	2	4	0.334
0	0	4	0.200
22	10	4	0.083
16	13	5	0.058
4	3	3	0.032
11	16	3	0.024
13	1	2	0.021
6	9	2	0.017
19	15	2	0.015
3	12	2	0.013
21	30	3	0.013
52	22	6	0.013
10	14	3	0.013
18	17	2	0.008
18	42	4	0.007
2	10	2	0.007
9	7	1	0.006
12	27	3	0.005
17	2	3	0.005
9	20	2	0.004
0	14	2	0.004
10	37	4	0.004
47	10	4	0.004
25	29	3	0.004
18	23	1	0.003

Table 2.2: Inferred alignment structures and recovery frequencies for the most frequently recovered products for noted substrates.

CHAPTER 3: ROLE OF TMEJ IN CHROMOSOMAL DOUBLE STRAND BREAK REPAIR

3.1 Introduction

Repair of chromosome double strand breaks (DSBs) is essential for normal cell growth and resistance to exogenous break-inducing agents. At the organismal level DSB repair also protects against cancer, shapes the response to cancer therapy, and has specialized roles in germ cell and immune system development. Mammalian cells typically rely on two pathways for DSB repair: classically defined non-homologous end joining (NHEJ) and homologous recombination (HR). Resection, the excision of 5' terminal strands of the broken ends to expose long 3' ssDNA tails, is an essential intermediate in HR and helps channel DSBs to repair by this pathway[67].

A third DSB repair pathway, termed Alternative NHEJ or microhomology-mediated end joining (Alt-NHEJ/MMEJ), has also been described. Alt-NHEJ was initially defined as both independent of factors required for NHEJ and associated with deletions clustered at 3 to 6 base pair (bp) sequence identities ("microhomologies") in flanking DNA[15,76]. Such products also typically require factors that promote resection (e.g. CtIP and Mre11)[126] indicating that, like HR, Alt-NHEJ typically acts downstream of this step.

Alt-NHEJ/MMEJ in metazoans has also been linked to activity of DNA polymerase theta (Pol θ; encoded by *Polq*)[107]. By comparison, an ortholog for metazoan Pol θ is missing in yeast[126], where synthesis during MMEJ has instead been associated with the

combined activity of multiple polymerases[80]. Pol θ is a 290kDa protein with a super family 2 helicase-like domain in the N-terminal third, and a Family A DNA polymerase domain in the C-terminal third[87]. Both *in vitro* and cellular experiments indicate it promotes MMEJ by using one 3' ssDNA overhang as a synthesis primer, after annealing this overhang to a second ssDNA overhang at a short tract of complementary sequence [107]. Strikingly, Pol θ is frequently over-expressed in HR-deficient cancers and overexpression is linked to poor prognosis[89,110,127]. Pol θ can also be essential in the context of HR deficient cancer cell lines[89,108] identifying it as a promising therapeutic target.

The role of Pol θ in cellular repair and the relationship between this pathway and canonical DSB repair pathways (NHEJ, HR) are still unclear. We investigated these questions through use of chromosomal assays and a panel of mouse embryo fibroblast lines (MEF) cell lines deficient in *Polq*, NHEJ (*Ku70*), and factors implicated in control of end resection and repair pathway choice (*Mre11, 53BP1*). Our results provide insight into how Pol θ activity contributes to cellular DSB repair. We show that Pol θ activity is essential when canonical repair pathways are impaired. We further identify an essential role for this enzyme in rescuing repair of DSBs in contexts where the ability to appropriately regulate end-resection has been compromised.

3.2 Methods

Cell lines

Experiments varying *Polq* status employed SV-40 Large T antigen transformed MEFs recovered from mice where a Polq null allele was generated by conventional gene targeting and made isogenic by backcrossing in mice. *Polq*$^{-/-}$ MEFs were then

complemented by lentiviral delivery of cDNA coding wild type human Polq, a polymerase domain variant harboring D2330A and Y2331A mutations, or a helicase domain variant harboring a K121M mutation**.** Deficiency or complementation was assessed by RT-qPCR for mRNA expression relative to a housekeeping control (GAPDH).

Variants of the above lines with frameshift mutations in Exon 3 of the *Ku70* gene were made following transient expression of Cas9 and sgRNA, and complemented variants of these Ku70 deficient lines made with lentivirus expressing mouse Ku70. Ku70 deficiency or complementation was validated by western analysis and a functional assay for NHEJ activity. In Figures 3.1 and Figure 3.3 experiments, the relevant genes were mutagenized with lentivirus that co-expressed Cas9 and sgRNA.

Chromosomal double strand break repair assays

To assess chromosomal end joining, plasmids expressing Cas9 (5 ug) and *Rosa26* sgRNA (5 ug) were introduced into $2x10^6$ cells by electroporation (Neon, Invitrogen) using a 1350 V, 30 ms pulse in a 100 µL chamber (Table 3.1). Cells were re-plated into DMEM without penicillin-streptomycin. Genomic DNA was purified at 6, 24, and 48 hours post-transfection (QiaAMP kit, Qiagen) and products were characterized by polyacrylamide gel electrophoresis on a 6% gel and by amplicon sequencing of a 472 bp region flanking the Cas9 cut site.

To assess gene targeting through homologous recombination at the *Rosa26* locus, plasmids expressing Cas9 (5 ug) and *Rosa26* sgRNA (5 ug), and a homology donor (20ug) that destroys the Cas9 recognition site by integration of a non-endogenous ScaI restriction enzyme site were introduced into $2x10^6$ cells by electroporation (Neon,

Invitrogen) using a 1350 V, 30 ms pulse in a 100 µL chamber. Cells were re-plated into DMEM without penicillin-streptomycin. Genomic DNA was purified 48 hours post-transfection (QiaAMP kit, Qiagen). Homologous recombination was measured by PCR amplification of the target site and digestion with ScaI, followed by polyacrylamide gel electrophoresis on a 5% gel of the target site and quantification of relative band intensities (ImageStudio). Dependency on Rad51 was confirmed by transfection of 20 nM control or Rad51 targeting siRNA pools prepared with RNAimax (Thermo). Depletion of Rad51 was confirmed by western blot.

To assess translocation formation, plasmids expressing Cas9 (5 ug) and *Rosa26* sgRNA (5 ug), and a plasmid expressing *H3f3b* sgRNA (5 ug) were introduced into $2x10^6$ cells by electroporation (Neon, Invitrogen) using a 1350 V, 30 ms pulse in a 100 µL chamber. Cells were re-plated into DMEM without penicillin-streptomycin. Genomic DNA was purified 48 hours post-transfection (QiaAMP kit, Qiagen). Translocation frequency was determined by digital droplet PCR (ddPCR) quantification. To detect translocations, 35 ng of genomic DNA was emulsified into ~20,000 independent oil-based droplets in the presence of Taq man reagents and 2x Master Mix for probes (BioRad) using a QX100 Droplet Generator (BioRad). Parallel reactions and reagents were used to score translocations (Translocation Fwd/Rev/Probe) and number of genomes sampled (Genome Ctrl Fwd/Rev/Probe). Droplets were amplified for 40 cycles, then read for fluorescence using a QX100 Droplet Reader (BioRad), and analyzed with Quantasoft ddPCR software. Parallel reactions were also chloroform extracted, and sequenced using standard methods.

Sequencing

Template DNA for each sequencing library was from independent electroporations performed across three days. $2x10^4$ input junction molecules were amplified using Phusion DNA polymerase (NEB) and PCR primers with 6 nt barcode sequences appended to their 5' ends for 22 cycles (Table 3.1). 15.5 ng of amplified DNA from each library was pooled in groups of ~10 libraries, phosphorylated with T4 PNK (NEB), treated with dATP and Klenow Exo – (NEB) to adenylate the 3' termini. Adaptors for sequencing (Illumina) were then appended to the amplicons by ligation, and free adaptor removed by gel purification. Gel purified pooled libraries were subjected to a 10 cycle enrichment amplification with adaptor-specific primers and purified with Ampure XP beads (Agencourt, Beckman Coulter). Equal amounts of DNA from each pooled library set were combined and submitted for a 2x300 paired end sequencing run (MiSeq, Illumina), with a PhiX174 DNA spike.

After sequencing, reads of PhiX174 DNA were removed. Remaining reads were then trimmed for quality, paired ends were merged, and libraries were de-indexed using Genomics workbench v8.0 (CLC-Bio). Greater than 50,000 reads/replicate were typically recovered for each replicate; reads recovered less frequently than expected given the oversampling ratio (typically about 5) were discarded as likely products of error during sample processing. Characterization of substrate repair junctions was performed by alignment of reads to a reference junction, export in SAM format, and deconstruction of the CIGAR string in Excel (Microsoft) to yield parameters including flanking deletion, length of ssDNA tails generated, and microhomology content at the junction. Deletion maps were generated using the ggplot package in R-Studio.

Insertions >24 bp (Figure 3.2, Table 3.2) were mapped against the reference mouse genome with batch BLAST searches.

We additionally assessed the extent to which proportional sampling of the population was maintained under the experimental conditions by amplifying and sequencing a control library of four cloned junctions of varying fragment size where the input number of template molecules was comparable to cellular experiments. The mean frequency of recovery for each of the four products was 25.4% (+/- 2.8%), and thus not significantly different than expected assuming equal efficiency of recovery (25%).

Cell cycle analysis

S-phase cells were quantified by flow cytometry after cells were treated with 10 uM BrdU (Sigma) for 1 hour, fixed, washed with PBS, treated with 0.1 M HCl/0.5% Triton X-100, and stained for both total DNA and incorporated BrdU.

Chromosomal aberrations

Chromosomal aberrations (breaks, fragments, and fusions) were counted after cells were treated with 50 ng/ml colcemid for 45 min, lysed, fixed, and stained with Giemsa (Gibco).

Immunofluorescence

Cells were irradiated with 1 Gy of gamma radiation and incubated for 6 hours, then washed, fixed, permeabilized, and stained to detect nuclear material (DAPI) and Rad51 foci. Images were acquired using a Pla-ApochroMAT, 63x /1.4 oil DIC objective lens (Zeiss Axio Imager 2, Duke University). Nuclei containing >10 foci were counted as positive, with 100 cells counted for each experiment.

Cell proliferation assays

To generate growth curves of the MEF panel described above, cells of each genotype were plated at $5x10^3$ cells per well in 24 well dishes. Cells were allowed to recover from plating for 12 hours, then were trypsinized, and counted on a hemocytometer every 24 hours over a 5-day period. Proliferation was quantified relative to the number of cells present for each condition at 12 hours post-plating.

To assess colony forming ability, Cas9 and sgRNAs targeting either control genes, or genes of interest were delivered to MEF cells via lentiviral transduction of a construct co-expressing both CRISPR/Cas9 components. 48 hours post-transduction, cells were plated at low density into 10 cm plates containing 2 ug/ml puromycin to select for transduced cells. Plates were fixed and stained 10 days later with 6% glutaraldehyde and 0.5% crystal violet. Colonies of >100 cells were counted.

Resection Assay

To measure the amount of resected DNA present at the chromosome 6 cut site we adapted a previously described assay for quantifying ssDNA in the genome by qPCR[128]. In short, PstI resection enzyme sites were chosen at intervals downstream (50 bp and 400 bp) of the chromosome 6 cut site. 150 ng of genomic DNA was digested overnight with PstI in a 20 ul reaction volume and used as template for a Taq man qPCR flanking the restriction enzyme site. Single stranded DNA is insensitive to PstI digestion and % ssDNA is quantified based on qPCR results.

3.3 Results

To better address the relationship between TMEJ and classically defined NHEJ we sought to make cell lines deficient in one or the other pathway, or both. We consequently employed CRISPR/Cas9 to generate variants of existing isogenic wild type and *Polq*$^{-/-}$ MEF lines[107] that do not express significant Ku70 protein. These lines are deficient in NHEJ, and can be complemented by expression of introduced Ku70 cDNA. Importantly, bi-allelic mutation of the *Ku70* gene in *Polq*$^{-/-}$ cells was rare, and relative to the other cell types the isolated *Polq*$^{-/-}$*Ku70*$^{-/-}$ cell line had both low colony forming efficiency and a severe proliferative defect, (Figure 3.1A). Combined loss of Ku70 and Pol θ is also associated with a reduced fraction of cells in S phase and increased chromosome aberrations (Figure 3.7A, 3.7B). Expression of a Ku70 cDNA in this line was sufficient to correct the growth defect, confirming it could be attributed to combined loss of Ku70 and Pol θ. To further address this issue we performed a reciprocal experiment, by mutagenizing Pol θ in the context of cells already deficient in *Ku70*. Lentiviral infection of *Ku70* deficient cells with 3 different constructs that target mutations to the *Polq* gene, but not 3 different control target sites, severely reduced both the frequency and size of colonies formed (Figure 3.1B). No significant differences in colony forming ability were observed when the same 6 viruses were introduced into wild type cells (Figure 3.7C). Deficiencies in Ku70 and Pol θ thus have synergistic effects on cell growth and viability - the doubly deficient cells are “synthetic sick”.

TMEJ rescues chromosome break repair when canonical pathways are compromised

We next addressed the relationship between TMEJ and canonical pathways (NHEJ, HR) in repair of a chromosome break by introducing a DSB with the Cas9 nuclease targeted to the *Rosa26* locus in Chromosome 6 (Figure 3.2A). We then determined the effect of Pol θ and Ku70 deficiencies on repair of this break by end joining (Figure 3.2, Figure 3.3) and HR (Figure 3.5), as well as the extent to which it is prone to participating in translocations with another Cas9-induced break (Figure 3.5). We confirmed all cell types expressed comparable amounts of Cas9 (Figure 3.8A), and assessed end joining by sequencing the region flanking the target site. As it is not possible to distinguish uncut chromosome from error free repair, we consider as definitive products of end joining only those sequences with insertions or deletions at the target site.

The majority (75%) of products in wild type (WT) cells had short 1-5 bp deletions, and deficiency in Pol θ had little impact on either the frequency or character of these products (Figure 3.2B, Figure 3.8B). This is consistent with a predominant role for NHEJ in repairing these breaks. By comparison, deletions that were both intermediate in size (between 5 and 50 bp) and associated with microhomologies were depleted in Pol θ deficient cells (reduced 4 fold; Figure 3.2C). Expression of a WT *Polq* cDNA was sufficient to reverse this defect (Figure 3.8C), and the effect of Pol θ on microhomology associated deletion was apparent for all microhomology lengths greater than 2 bp (Figure 3.8D).

Deficiency in Ku70/NHEJ resulted in deletions that were both much larger and more frequently associated with microhomology (Figure 3.2B, 3.2C). These microhomology-

associated deletions were again mostly dependent on Pol θ (reduced 7 fold in *Polq*$^{-/-}$ *Ku70*$^{-/-}$ cells), but the effect of Pol θ was no longer limited to the subset of deletions ranging from 5-50 bp (Figure 3.2C). We additionally confirmed expression of a Ku70 cDNA in both Ku70 deficient cells (*Ku70*$^{-/-}$, *Polq*$^{-/-}$*Ku70*$^{-/-}$) was sufficient to return the frequency and pattern of microhomology associated deletions to that observed in the matched Ku70$^{+/+}$ parent cell lines (Figure 3.8C). Comparison of end joining spectra over time also indicates that TMEJ is engaged much more rapidly in Ku deficient cells (Figure 3.8E). Therefore, relative to an NHEJ proficient background, TMEJ in Ku70 deficient cells accounts for a much larger fraction of repair, is engaged faster, and its influence now extends to large deletions.

Two classes of insertions had conspicuously distinct genetic requirements. Most insertions were short (1-4 bp) and generated during repair by NHEJ (reduced 7 fold in *Ku70*$^{-/-}$ cells; Table 3.2). In contrast, deficiency in Ku70 alone had the opposite effect on longer insertions (3 fold increase), and these longer insertions were dependent on Pol θ (Figure 3.2D). To assess if these Pol θ-dependent insertions were templated, we focused on the subset of insertions that were sufficiently large (>24 bp) to allow mapping to the mouse genome. Most of these insertions – 87% of those observed in *Ku70*$^{-/-}$ cells – could be mapped (i.e. were definitively templated) to sequence near the Cas9 target site in chromosome 6 (Table 3.2). A surprisingly high frequency ($1.5X10^{-3}$ in *Ku70*$^{-/-}$ cells) of end joining products recovered had insertions that used template from other chromosomes (Table 3.2), and were derived from both genic and intergenic regions. Template switching events, both intra- and inter-chromosomal, were often complex (Figure 3.2E), with successive segments derived from template tracts

separated by 100s of base pairs, and 2-3 bp microhomologies often evident in the flanks of successive segments. In some cases successive segments were derived from nearby DNA but employed the opposite strand as template. Insertions mediated by Pol θ are thus largely template-dependent, and provide striking evidence of the ability of this polymerase to promiscuously switch template.

We show above that in Ku deficient cells TMEJ/ Pol θ is both engaged more rapidly and accounts for a much greater fraction of repair. Since extrachromosomal experiments indicated Pol θ acts only on pre-resected ends, we reasoned the synthetic effects of Pol θ and Ku deficiency on proliferative capacity could reflect excessive end resection associated with Ku70 deficiency, and a central role for TMEJ in rescuing repair of these ends. In accord with this model, mutagenesis of the *Mre11* gene, a factor implicated in end resection, had an impact on large deletion in *Ku70*$^{-/-}$ cells comparable to loss of Pol θ (Figure 3.3A). Direct measurement of resected DNA surrounding the chromosome 6 break, quantified by PstI restriction endonuclease sensitivity, showed that resection is significantly increased in Ku70-deficienct cells, and that resection increases over time (Figure 3.4A). This increase in resection is fully ablated upon compound mutagenesis of the *Ku70* and *Mre11* genes (Figure 3.4B). This model also predicts there should be a genetic interaction between *Polq* and *53BP1*, a factor that promotes NHEJ in part by antagonizing end resection[59,129]. Strikingly, lentiviral constructs engineered to mutate the *53BP1* gene, but not off-target controls, reduced both the frequency and size of colonies formed in a Pol θ deficient background (Figure 3.3B, Figure 3.9). By comparison, there was no difference comparing targeting vs. non-targeting constructs in parallel experiments in wild type cells (Figure 3.9). Our results

thus identify a central role for TMEJ in repairing chromosome breaks when resection is excessive or mis-regulated.

We next sought to address whether defects in TMEJ might influence the efficiency of repair by HR of the same break studied in the experiments described above (Figure 3.2 and Figure 3.3A). We generated breaks at the Cas9 site in chromosome 6 as above, but now included an extrachromosomal homology donor with 6 bp substituted at the target site that both ablates Cas9 recognition and introduces a site for the restriction endonuclease ScaI (Figure 3.5A). Products of HR were thus scored as ScaI sensitive species in amplified products of the native chromosome flanking the Cas9 target; these were sensitive to depletion of Rad51 (Figure 3.10A) and dependent on both the presence of the homology donor and the initiating DSB (Figure 3.5B). We observed a frequency of HR in WT MEFs of 10%, and this is increased more than 2 fold in Ku70 deficient cells (Figure 3.5C, Figure 3.10B). By comparison, deficiency in Pol θ did not similarly result in increased HR. Indeed, loss of Pol θ in cells deficient in Ku70, where TMEJ otherwise accounts for a large fraction of repair, led instead to a slight decrease in HR efficiency. We note loss of Pol θ does increase a marker of an intermediate step in HR, ionizing radiation-induced Rad51 foci[89,108], but this effect was again less than that attributable to deficiency in Ku70 (Figure 3.10C, 3.10D). We conclude NHEJ effectively competes with HR for repair of nuclease-induced breaks (in accord with previous work)[130], but that TMEJ does not.

Simultaneous breakage at two sites in the genome can lead to the joining of ends from different chromosomes, or translocation. We addressed the contribution of TMEJ to this process by again introducing breaks at the previously used Cas9 site in

chromosome 6, but now also introduced a second break at a site in chromosome 11 (Figure 3.5D). Genomic DNA was then harvested and used as template in emulsion PCRs (ddPCRs) to assess the frequency of the t(6;11) der11 translocation product. In accord with previous work[131], translocation frequency relative to WT cells was increased 3 fold in Ku70 deficient cells (Figure 3.5E). However, deficiency in Pol θ alone had no significant impact. Moreover, characterization of translocation products indicated that for each of the four cell types (WT, *Polq*$^{-/-}$, *Ku70*$^{-/-}$, and *Polq*$^{-/-}$*Ku70*$^{-/-}$), the frequency of junctions with microhomology was reduced relative to the products of intra-chromosomal deletion (Table 3.3). TMEJ thus does not promote formation of translocations under the conditions used here. More significantly, the frequency of translocation was highest when cells were deficient in both Pol θ and Ku70. TMEJ can thus protect against this class of chromosome rearrangement, though this effect was significant here only if cells are unable to efficiently repair breaks by NHEJ.

3.4 Discussion

Our results indicate that in cells, Pol θ has a crucial role as a back up pathway to canonical double strand break repair mechanisms. We show deficiency in Pol θ alone has at best minor effects on activities of other pathways (NHEJ and HR) in the repair of a Cas9-induced break (Figure 3.2, Figure 3.5). TMEJ thus does not typically compete with these pathways. However, TMEJ has a critical role when there is excessive resection, i.e. cells deficient in Ku70 or 53BP1 (Figure 3.1, Figure 3.2, Figure 3.3). We conclude Pol θ/TMEJ is essential for repair of DSBs when either major DSB repair

pathway is compromised (Figure 3.6), and both protects against or promotes genome instability dependent on context (Figure 3.5 and discussed below).

Role of TMEJ in double strand break repair

What is the relationship between previously defined Alt-NHEJ/MMEJ and Pol θ-dependent repair/TMEJ? Microhomology associated deletions typically used to define Alt-NHEJ/MMEJ are severely depleted in both Pol θ deficient contexts (i.e. *Polq*$^{-/-}$ and *Polq*$^{-/-}$*Ku70*$^{-/-}$), relative to their matched *Polq* proficient counterparts (Figure 3.2). Similarly, Alt-NHEJ has been defined as dependent on prior resection of ends, and we determined the effect of Pol θ deficiency on microhomology associated deletions is comparable to deficiency in a factor required for end resection, Mre11 (Figure 3.3A). There is also little evidence for repair of pre-resected extrachromosomal substrates in cells deficient in both Pol θ and Ku70 (Chapter 2, Figure 2.2). Taken together, our evidence indicates TMEJ accounts for the majority of events previously defined as MMEJ or Alt-NHEJ. Nevertheless, chromosomal end joining products still accumulate in *Polq*$^{-/-}$*Ku70*$^{-/-}$ cells (Figure 3.2B). This may represent residual repair mediated by the remaining NHEJ components, or an as yet un-defined and largely microhomology-independent alternative to the end joining pathways referenced so far (Alt-NHEJ/MMEJ/TMEJ and NHEJ).

In cells proficient in both NHEJ and HR, loss of Pol θ significantly affects only a small fraction (<5%) of DSB repair (Figure 3.2C), consistent with mild phenotypes of cells and mice deficient in Pol θ alone[98]. TMEJ is also associated only with deletions of intermediate size (5-50 bp) in this context. By comparison, TMEJ is engaged more rapidly, more often, and affects a wider range (5-250 bp) of deletions in cells deficient in

Ku70/NHEJ. Cell growth and viability are also severely impaired in cells deficient in both Pol θ and Ku70 (Figure 3.1), or Pol θ and 53BP1 (Figure 3.3); the latter factor is more specifically associated with attenuation of end resection than is Ku70. Our results are thus consistent with a central role for TMEJ when end structures are poor substrates for NHEJ, as well as in cells prone to excessive end resection (Figure 3.6).

Homology-dependent repair of Cas9-induced breaks was stimulated more than two fold by loss of Ku70. In contrast, loss of Pol θ had no significant effect on HR of the same break, even in a Ku70 deficient background. TMEJ is thus a less effective competitor than is NHEJ for repair of DSBs, at least when the DSBs are destined for successful repair. A similar lack of effect of Pol θ deficiency on repair by HR is evident in *C. elegans*[102]. Pol θ loss is nevertheless sufficient to stimulate IR-dependent Rad51 foci[89,108] (Figure 3.10C, 3.10D), which could reflect the accumulation of products of aborted HR that are no longer resolved by Pol θ-dependent means.

Previous work identified a key role for TMEJ in repair of chromosome breaks in homologous-recombination deficient cancer cell lines[89,108]. Here we find the events we can attribute to Pol θ/TMEJ – deletions that are less than 50 bp and associated with flanking microhomologies >2 bp or insertions > 4 bp – is precisely equivalent to a mutational signature associated with those breast, ovarian, and pancreatic cancers that have mutations in BRCA1 and BRCA2 (COSMIC Signature 3)[132].

Strikingly, 53BP1 loss has opposite effects on cell viability when comparing cells deficient in Pol θ vs. cells deficient in BRCA1; it impairs growth and viability in the former context (Figure 3.3B), but rescues it in the latter[59,129]. Of special note is the ability of 53BP1 loss to “protect” BRCA1 deficient cancers from therapies that employ

inhibitors of poly(ADP)ribose polymerase. Our data indicate Pol θ/TMEJ activity will be essential in the context of this mechanism of therapy resistance. These results emphasize the critical importance of appropriately regulated end-resection, and further emphasizes the importance of Pol θ/TMEJ as a therapeutic target.

TMEJ and genome instability

There is evidence that Pol θ/TMEJ both suppresses and promotes DSB-induced chromosome rearrangement. Pol θ was identified in a systematic screen for suppressors of spontaneous genome instability[97], and it additionally suppresses translocations associated with immunoglobulin class switch recombination[107]. We report here that Pol θ can also suppress translocations between a pair of Cas9-induced breaks targeted to different chromosomes, though a significant effect was limited to cells already deficient in NHEJ (Figure 3.5B). Moreover, the fraction of these translocation products associated with a TMEJ-like signature (e.g. microhomologies) was reduced relative to matched products of intra-chromosomal deletion for all four cell types.

In contrast, a previous report assessing Cas9-induced translocations indicated these were promoted by Pol θ, and Pol θ was required in NHEJ deficient cells for fusion of chromosomes at de-protected telomeres[108]. Additionally, this and related studies[133] indicate translocation products are often enriched for TMEJ-associated signatures. Moreover, as also noted above, this signature is associated with deletions observed in hereditary breast cancer[132].

Results described here also suggest how Pol θ/TMEJ could promote chromosome rearrangement. Analysis of intra-chromosomal deletions identified frequent “template-switching” events; these insertions were assembled from multiple template segments,

sometimes from opposite strands, and often with 2-4 bp microhomologies associated with the flanks of successive segments. Importantly, promiscuity in template switching extended to surprisingly frequent switching to template tracts in other chromosomes ($1.5x10^{-3}$ of total products in Ku70 deficient cells). We note our ability to recover examples of switching of template to another chromosome required a "switching back" of the repair process, such that the rejoined product remained intrachromosomal. It thus seems likely there is an undetected fraction of these events – perhaps even the majority – that never switch back, and contribute to chromosome aberrations similar to the microhomology-mediated break induced replication (MM-BIR) events described in yeast and mediated by Pol ζ and Rev1[134].

This varied contribution of Pol θ/TMEJ to genome instability in mammalian cells may reflect differences in the origin of the chromosome break and cell type. Notably, evidence that TMEJ suppresses genome instability comes from analysis of populations of primary cells, and where the assessed chromosome aberrations arise from chromosome breakage that was spontaneous[98] or naturally programmed[107]. The properly regulated use of TMEJ as a backup mechanism thus may protect against genome instability. However, as is apparent in hereditary breast cancer, defective DNA damage responses in transformed cell lines and cancers can result in excessive engagement of TMEJ and even an “addiction” to this pathway[89,108,110,127]. The ability of Pol θ/TMEJ to promote genome instability may thus be acquired in parallel with the transition of its role from backup to essential. In sum, this work identifies Pol θ as a critical factor in the survival and evolution of cancers deficient in canonical repair pathways, as well as a promising target for the treatment of these cancers.

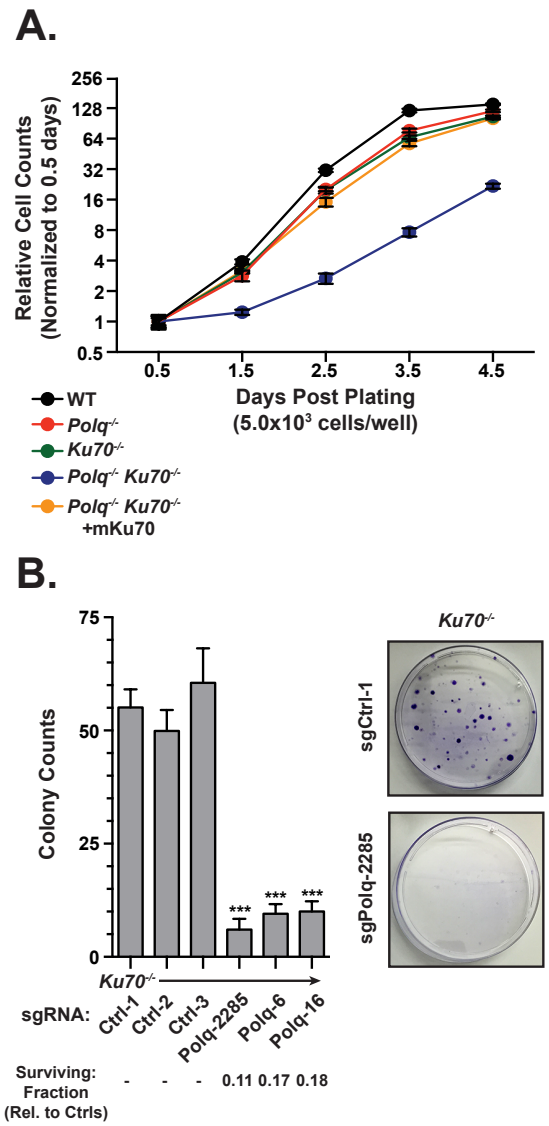

Figure 3.1: Effect of end joining deficiencies on cell proliferation

(A) The noted cell types were seeded at 5x10^3 cells/well in 24 well dishes, then counted 12 hours after plating and every 24 hours afterwards. Data shown are the mean +/- SEM, n=3. **(B)** *Ku70*$^{-/-}$ cells were infected with lentivirus containing Cas9 and either control guides or guides targeting sites in regions coding for mouse Pol θ, seeded into 10 cm plates, and stained after 10 days (right panel). Colonies of more than 100 cells were counted from triplicate experiments (left panel); Data shown are the mean +/- SEM, n=3. Statistical significance was assessed by ANOVA with Bonferoni correction for multiple comparisons. ***p<0.001.

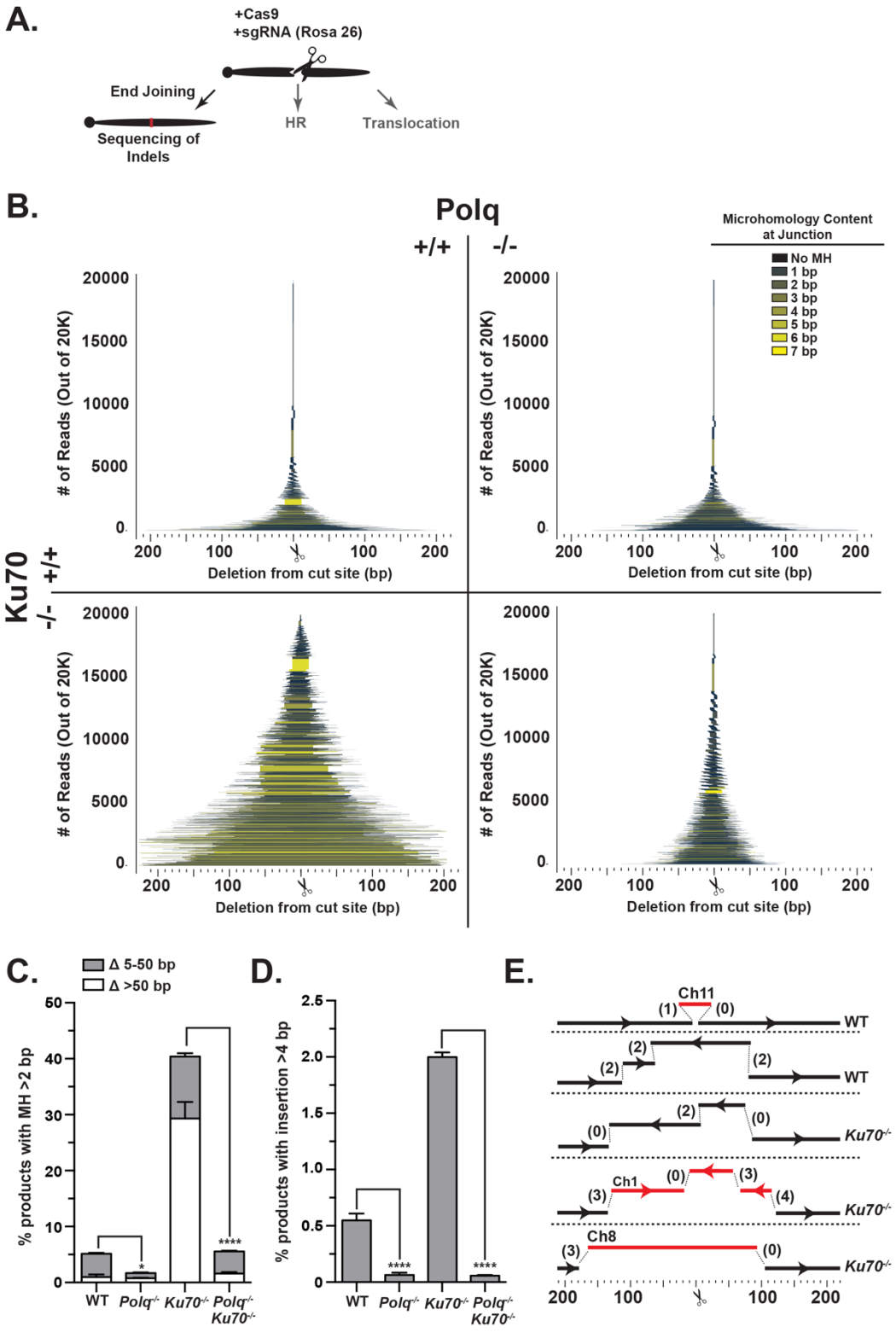

Figure 3.2: Effect of end joining deficiencies on repair of Cas9-induced chromosome breaks

Figure 3.2: Effect of end joining deficiencies on repair of Cas9-induced chromosome breaks

(A) Cas9 and guide RNA specific to a chromosome 6 target were expressed in the noted cell types, and products of end joining characterized by target site amplification and sequencing. **(B)** Bars denote extent of deleted DNA for each end joining product relative to the Cas9 target site, and are shaded according to the extent of microhomology as noted in the legend. The height of each bar on the Y axis defines the proportion of each product in a set of 20,000 estimated input molecules, and is averaged from 3 experiments for each of WT (top left), *Polq*$^{-/-}$ (top right), *Ku70*$^{-/-}$, (bottom left), and *Polq*$^{-/-}$*Ku70*$^{-/-}$ cells (bottom right). **(C)** The percentage of all end joining products with microhomology >2 bp and deletion (Δ) within the noted size ranges. Data shown are the mean +/- SEM, n=3. Statistical significance was assessed by two-way ANOVA with Bonferoni correction (results for the 5-50 bp category shown). **p*<0.05, *****p*<0.0001. **(D)** The percentage of all end joining products with insertion >4 bp recovered from MEF lines described above. Data shown are the mean +/- SEM, n=3. Statistical significance was assessed by one-way ANOVA with Bonferoni correction (*p* values as above). **(E)** Structure of selected Pol θ mediated insertions. Successive segments denote template switching, with opposing arrows identifying use of opposite strands. The lengths of microhomologies in flanks of successive segments are in parentheses. Black segments, from chromosome 6; red segments, from other chromosomes.

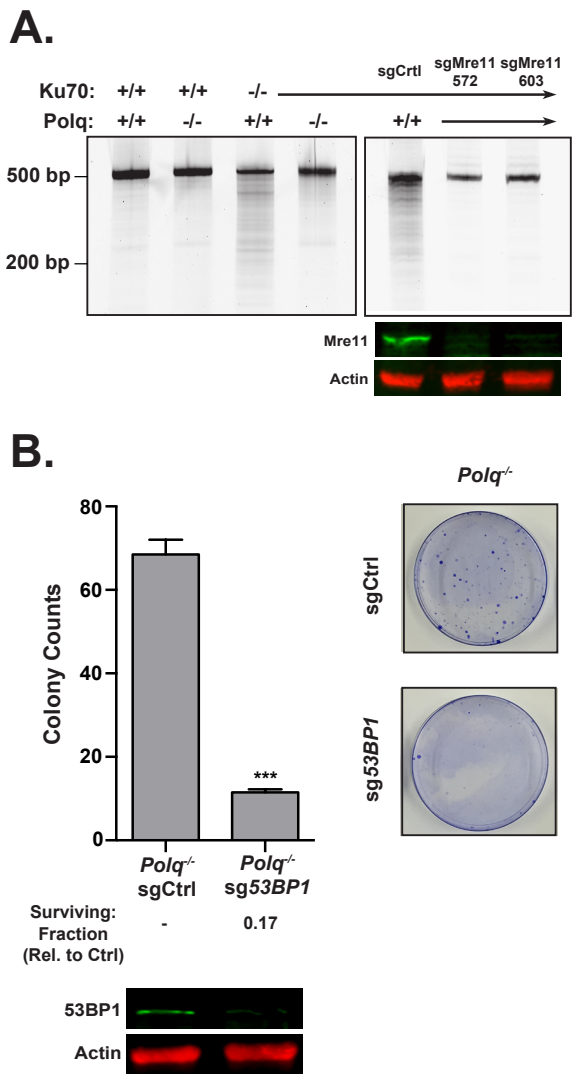

Figure 3.3: Effect of aberrant end resection on repair and viability

(A) Template DNA from the experiment described in Fig. 4 (left panel), or from *Ku70*$^{-/-}$ cells infected with lentivirus containing Cas9 and either a pool of control guides or guides targeting Mre11 at amino acids 572 or 603 respectively (right panel), was amplified and characterized by electrophoresis. **(B)** *Polq*$^{-/-}$ cells were infected with lentivirus containing Cas9 and a pool of either control guides or guides targeting *53BP1*. Colony forming ability was assessed as in Fig. 1B. Data shown are the mean +/- SEM, n=3. Statistical significance was assessed by unpaired T-test. ***p<0.001. See also Figure S5.

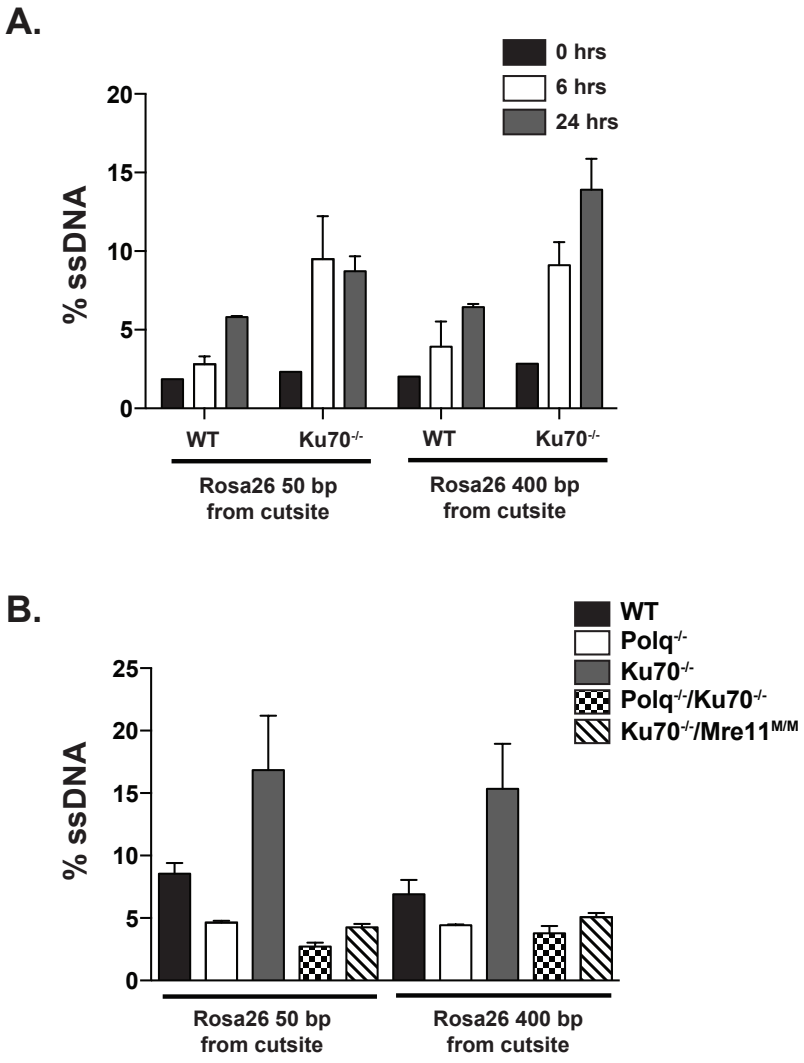

Figure 3.4: Direct measurement of resected DSBs

(A) Template DNA from the experiment described in Fig. 4 (left panel), or from *Ku70*$^{-/-}$ cells infected with lentivirus containing Cas9 and either a pool of control guides or guides targeting Mre11 at amino acids 572 or 603 respectively (right panel), was amplified and characterized by electrophoresis. **(B)** *Polq*$^{-/-}$ cells were infected with lentivirus containing Cas9 and a pool of either control guides or guides targeting *53BP1*. Colony forming ability was assessed as in Fig. 1B. Data shown are the mean +/- SEM, n=3. Statistical significance was assessed by unpaired T-test. ***p<0.001. See also Figure S5.

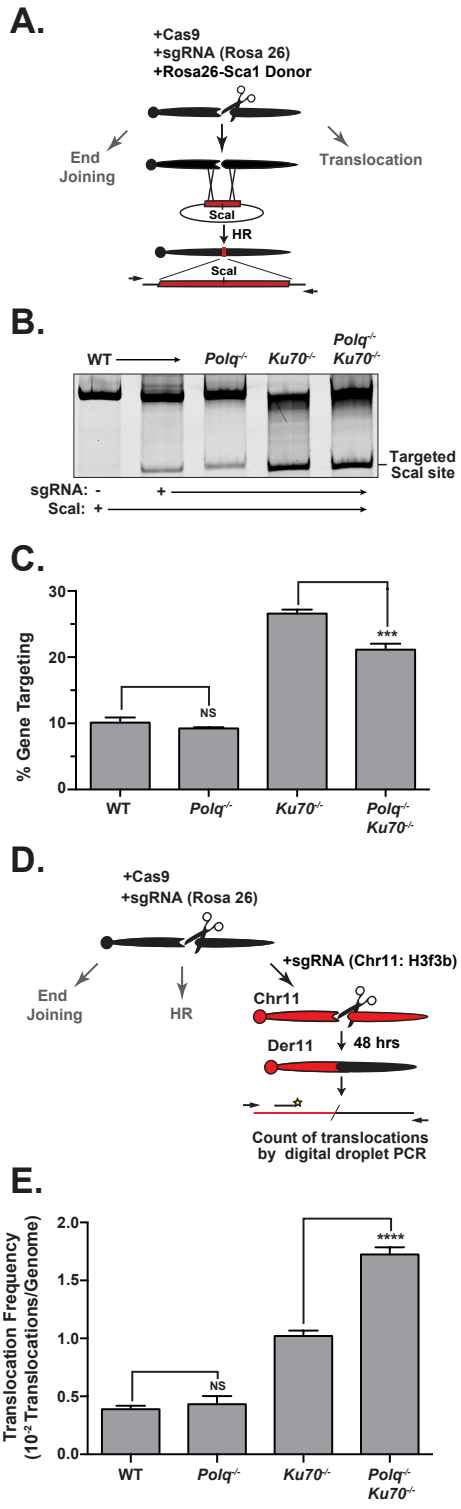

Figure 3.5: Effect of end joining deficiencies on repair of chromosomal breaks by homologous recombination and translocation

Figure 3.5: Effect of end joining deficiencies on repair of chromosomal breaks by homologous recombination and translocation

(A, B) Homologous recombination was assessed by introduction of Cas9, guide RNA, and a plasmid homology donor that introduces a ScaI recognition site at the site of breakage in chromosome 6 into noted cell types. DNA was harvested 48 hours later, amplified, and digested with ScaI followed by electrophoresis. **(C)** The mean fraction of the chromosome 6 target site that acquired sensitivity to ScaI. Data shown are the mean +/- SEM, n=3. Statistical significance assessed by one-way ANOVA with Bonferoni correction. NS =not significant, ***p<0.001, ****p<0.0001. **(D)** Cas9 and guide RNAs targeting chromosome 6 and chromosome 11 were expressed in the noted cell types. Genomic DNA was isolated after 48 hours and used as template for parallel emulsion PCRs (ddPCR) specific for the t(6,11) der11 translocation product and an input genome control. **(E)** Mean frequency of der11 translocations is determined by the number of translocations over the number of input genomes; Data shown are the mean +/- SEM, n=3, with >$4x10^4$ genomes assessed/experiment. Statistical significance was assessed by one-way ANOVA with Bonferoni correction (p values as above).

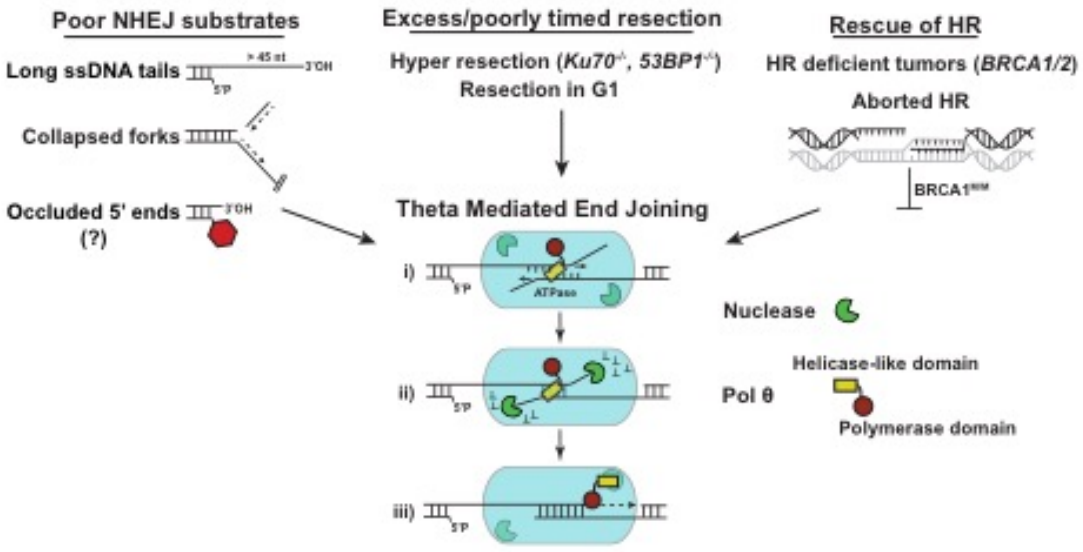

Figure 3.6: Mechanism and cellular roles of Pol θ-mediated end joining

Suggested contexts for engagement of Pol θ mediated end joining, and proposed mechanism including a i) search for microhomology, ii) removal of non-homologous tails, and iii) synthesis.

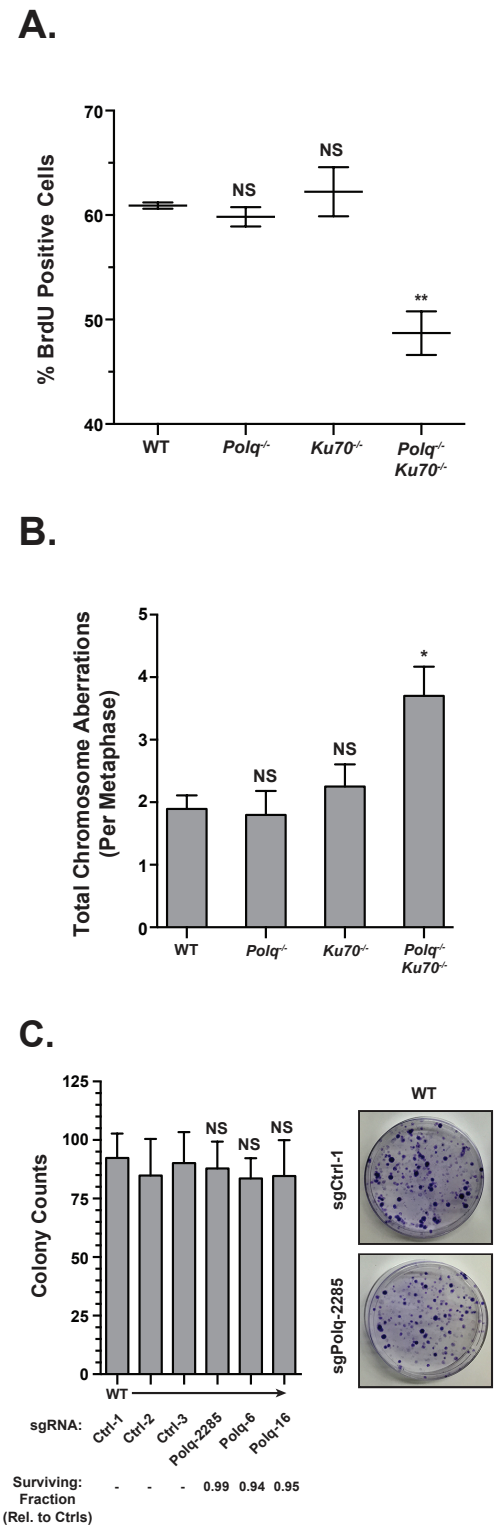

Figure 3.7: Characterization of Ku70/Pol θ synthetic sickness

Figure 3.7: Characterization of Ku70/Pol θ synthetic sickness

(A) BrdU incorporation was assessed in the noted cell types by flow cytometry. Data are the mean +/- SEM, n=3. Statistical significance assessed by one-way ANOVA with Bonferoni correction. **(B)** Total chromosome aberrations (chromatid breakage, fragments/small chromatin, and chromosome exchanges/fusions) were scored on metaphase spreads of the noted cell types. Data are the mean +/- SEM, n=3. Statistical significance assessed by one-way ANOVA with Bonferoni correction. **(C)** Wild type cells were infected with lentivirus containing Cas9 and either control guides or guides targeting sites in regions coding for mouse Pol θ, seeded into 10 cm plates, and stained after 10 days (right panel). Colonies of more than 100 cells were counted from triplicate experiments (left panel); Data shown are the mean +/- SEM, n=3.

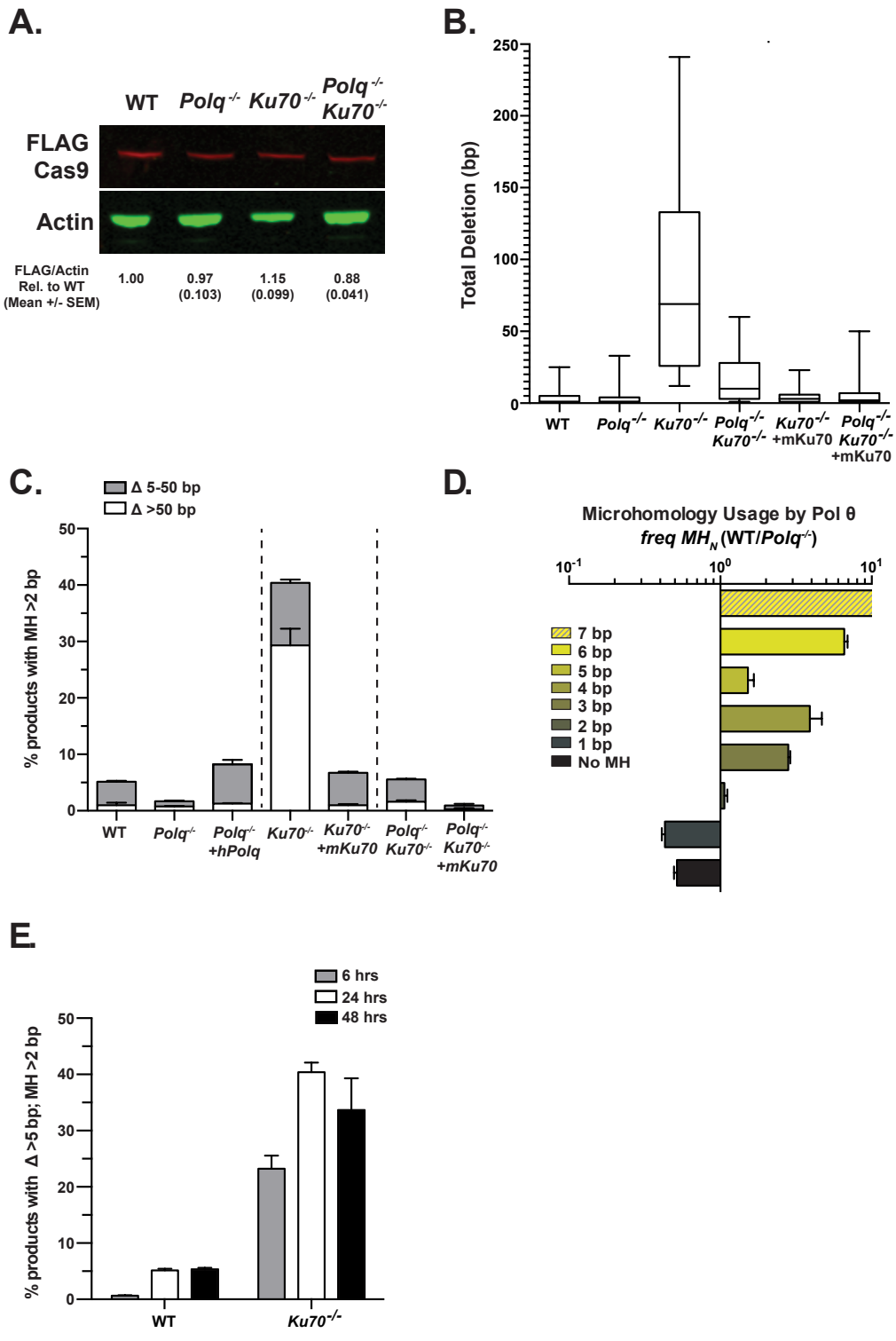

Figure 3.8: Supplemental characterization of effects of end joining deficiencies on repair of Cas9-induced chromosome breaks

Figure 3.8: Supplemental characterization of effects of end joining deficiencies on repair of Cas9-induced chromosome breaks

(A) Representative Western blot showing Cas9 expression 24 hours post transfection in noted cell types. Noted below are the mean levels of Cas9 expressed (normalized to correct for loading differences using the Actin control) in replicate transfections, relative to WT MEFs. **(B)** Box plots describe the inter-quartile range of deletions observed in the noted cell types. Whiskers describe deletions in the 10^{th} to 90^{th} percentile range.**(C)** The percentage of all end joining products with microhomology >2 nt and associated with deletion either between 5 and 50 bp or >50 bp recovered from noted cell types. Data are the mean +/- SEM, n=3. **(D)** The extent each sized microhomology is enriched (bars right of Y axis) or depleted (bars left of Y axis) is expressed as the fraction of products recovered from WT cells over the fraction of products recovered from *Polq*$^{-/-}$ cells for each category. **(E)** The percentage of all end joining products with microhomology >2 nt recovered from sequencing of DNA flanking the chromosome 6 Cas9 target site over 6, 24, and 48 hrs in the noted cell types. Data are the mean +/- SEM, n=3.

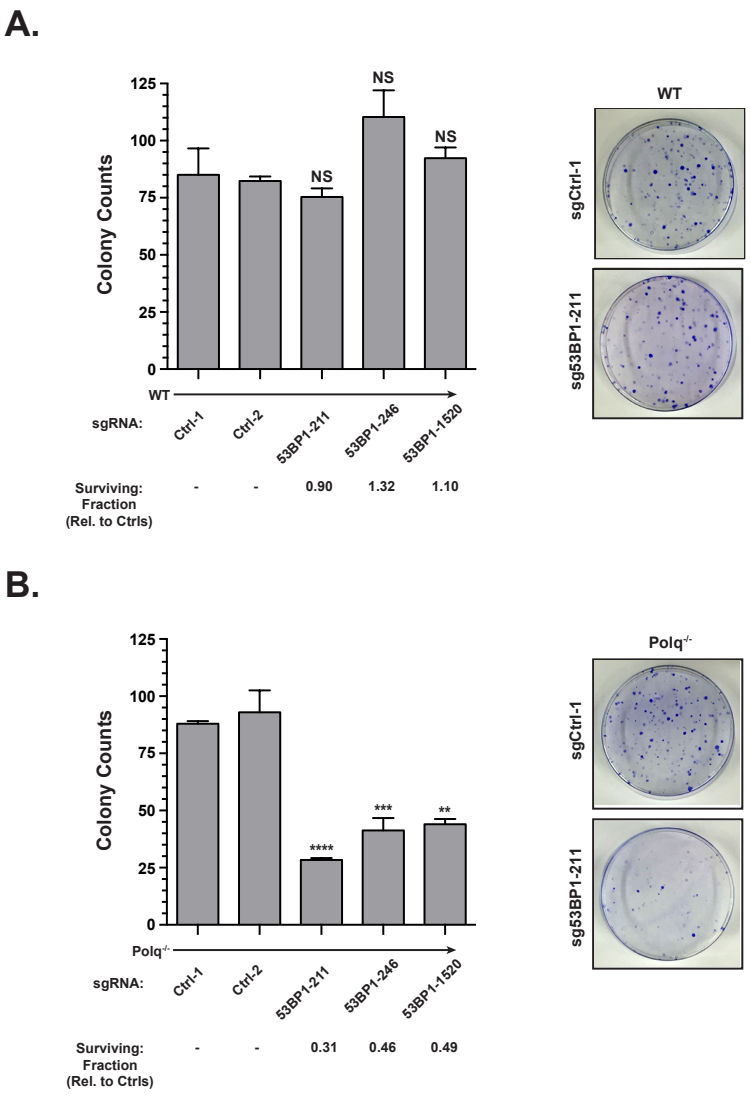

Figure 3.9: Effect of 53BP1 mutagenesis in WT cells

(A) WT cells were infected with lentivirus containing Cas9 and either control guides or guides targeting sites in regions coding for mouse 53BP1, seeded into 10 cm plates, and stained after 10 days (right panel). Colonies of more than 100 cells were counted from triplicate experiments (left panel); Data shown are the mean +/- SEM, n=3. Statistical significance assessed by one-way ANOVA with Bonferoni correction. **(B)** *Polq*$^{-/-}$ cells were infected with lentivirus containing Cas9 and either control guides or guides targeting sites in regions coding for mouse 53BP1, seeded into 10 cm plates, and stained after 10 days (right panel). Colonies of more than 100 cells were counted from triplicate experiments (left panel); Data shown are the mean +/- SEM, n=3. Statistical significance assessed by one-way ANOVA with Bonferoni correction.

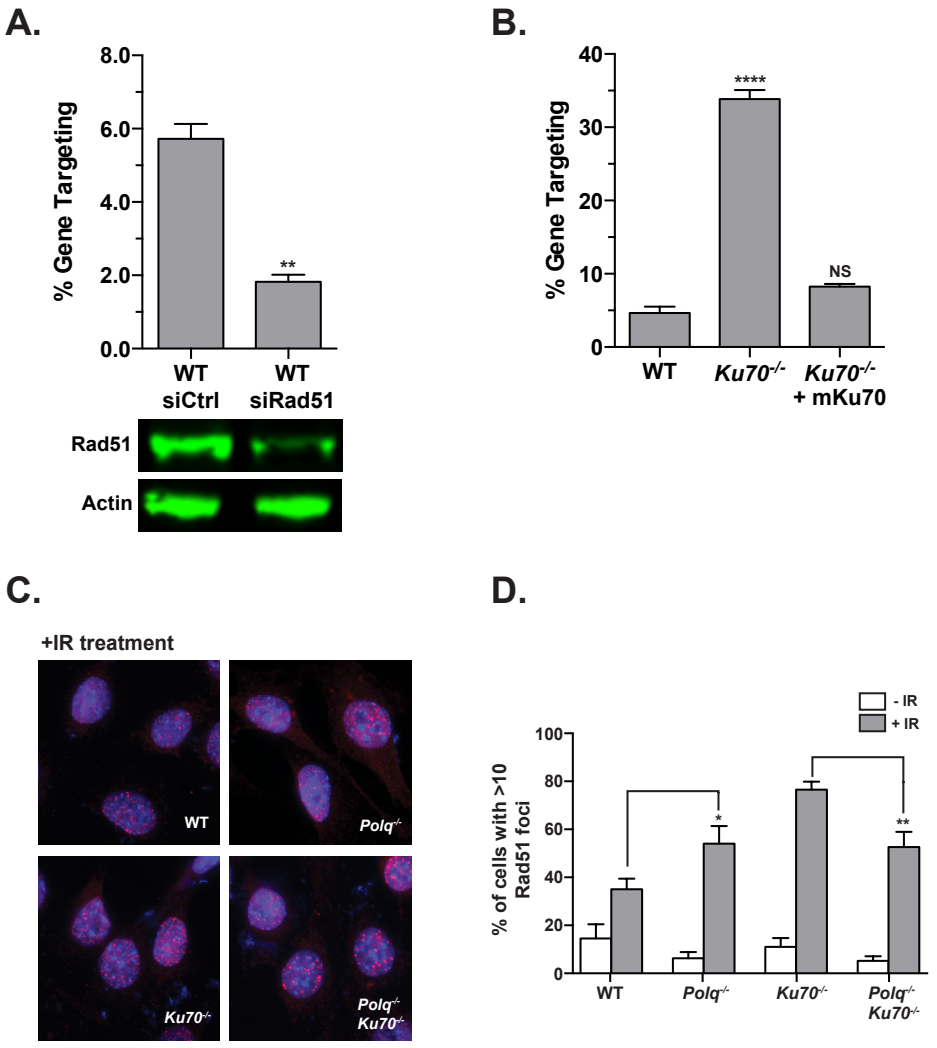

Figure 3.10: Effect of Rad51 depletion and Ku deficiency on homologous recombination

(A) Homologous recombination in WT MEFs transfected with either control or Rad51 siRNA pools, quantified as described in Figure 6. Data are the mean +/- SEM, n=3. Statistical significance assessed by one-way ANOVA with Bonferoni correction. **(B)** Homologous recombination in the noted cell types, quantified as described in Figure 6. Error bars represent the standard error of the mean of three independent electroporations. Data are the mean +/- SEM, n=3. Statistical significance assessed one-way ANOVA with Bonferoni correction. **(C)** Noted cell types were treated with 1 Gy of gamma radiation and co-stained for Rad51 foci (red) and nuclear material (DAPI, blue). **(D)** Nuclei from experiment in panel C containing >10 Rad51 foci were counted as positive, with 100 cells scored for duplicate experiments. Data are the mean +/- SEM, n=3. Statistical significance assessed by two-way ANOVA with Bonferoni correction (results for the +IR category are shown).

Primers/Probes	
Chromosome 6 Fwd	5'TCAGTTGGGCTGTTTTGGAG
Chromosome 6 Rev	5'GGCGGATCACAAGCAATAAT
HR Fwd	5'AGAATGCAGTGTTGAGGCC
HR Rev	5'AGAAAACTGGCCCTTGCCATT
Translocation Fwd	5'AGCCACAGTGCTCACATCAC
Translocation Rev	5'TCCCAAAGTCGCTCTGAGTT
Translocation Probe	5'/56-FAM/TCAGTAAGG/ZEN/GAGCTGCAGTGGAGTA/3IABkFQ/
Genome Ctrl Fwd	5'GGGAAGTGAGAGAGAAACTGAAG
Genome Ctrl Rev	5'AAACCTGAGCCAGACTTTCC
Genome Ctrl Probe	5'/5HEX/TCAGCAAAG/ZEN/ACCGCGGAAAGATCT/3IABkFQ/

sgRNA Sequences	
Ctrl-1	5'ACTCCAGTCTTTCTAGAAGA (Fig. 1)
Ctrl-2	5'CGCCGTGCACTGGTTCCGAA (Fig. 1)
Ctrl-3	5'GGCATCATTATTCAAATCAG (Fig. 1)
Polq-2285	5'CAGTTCAAGCTGAGAGTAGT (Fig. 1)
Polq-6	5'CGCCGTTTCCTACTCCGGCG (Fig. 1)
Polq-16	5'TCCCGAGAACGTGTCGGAGC (Fig. 1)
Chr 6 (Rosa 26)	5'ACTCCAGTCTTTCTAGAAGA (Fig. 4-6)
Chr 11 (H3f3b)	5'GTTGGCTCGCCGGATACGGG (Fig. 6)
Mre11-572	5'GGCCGAGGCCGAGGGCGAAG (Fig. 5)
Mre11-603	5'TGGAGATCACTACTCGAGGC (Fig. 5)
53BP1-211	5'TTCTAGCCCGCTATCTGATG (Fig. 5, S5)
53BP1-246	5'GTTACAGTACAGCCCGGTAA (Fig. 5, S5)
53BP1-1520	5'ACAAGCTGCTCTTTGATGAT (Fig. S5)

Antibodies	
FLAG Tag	Sigma F1804
53BP1	Bethyl Laboratories A300-272A
MRE11	Gift from John Petrini
Actin	Sigma A2066
Rad51	Novus nb100-48

SiRNA	
siCtrl Pool	Dharmacon D-001206-13-05
SiRad51 Pool	Dharmacon M-062730-01-0005

Plasmids	
FLAG-Cas9	A gift of Xingxu Huang (Addgene plasmid # 44758)
sgRNA Plasmid	A gift of Xingxu Huang (Addgene plasmid # 51133)
Lenti-CRISPRv2	A gift from Feng Zhang (Addgene plasmid # 52961)
Rosa 26 Homology Donor	605 bp of identity to sequence upstream of the Rosa26 target, a central segment sufficient to destroy recognition by the Rosa26 targeting guide and create a ScaI restriction enzyme site, followed by 537 bp of identity to sequence downstream of the Rosa 26 cut site, cloned into PCR2.1

Table 3.1: Reagents used during experiments.

Genotype	1-4 bp Short (SEM)	>4 bp Long (SEM)	>24 bp Very long "mappable" (SEM)		
			Ch 6	Other	Not mapped
WT	0.28 (0.02)	$5.5(0.6)\times10^{-3}$	$1.1(.02)\times10^{-3}$	$6.1(2.0)\times10^{-5}$	$4.3(0.3)\times10^{-4}$
Polq$^{-/-}$	0.21 (0.01)	$6.4(2.0)\times10^{-4}$	$6.9(3.5)\times10^{-5}$	0	$2.0(0.8)\times10^{-4}$
Ku70$^{-/-}$	0.04 (0.002)	$2.0(0.04)\times10^{-2}$	$5.3(0.9)\times10^{-3}$	$1.5(0.4)\times10^{-3}$	$1.0(0.7)\times10^{-3}$
Polq$^{-/-}$*Ku70*$^{-/-}$	0.06 (0.01)	$5.6(0.6)\times10^{-4}$	0(0)	$1.1(0.8)\times10^{-4}$	0

Table 3.2: Frequency of insertions observed in end joining products.

	Frequency of products with >2 bp microhomology	
Genotype	**Ch6 Intrachromosomal**	**t(6;11) translocation**
WT	(0.162)	1/24 (0.042)
Polq$^{-/-}$	(0.124)	0/22 (0)
Ku70$^{-/-}$	(0.399)	5/31 (0.161)
Polq$^{-/-}$ *Ku70*$^{-/-}$	(0.166)	2/21 (0.096)

Table 3.3: Frequency of products with >2 bp of microhomology.

CHAPTER 4: DISCUSSION

DNA double strand breaks are a structurally diverse and highly genotoxic class of damage that can result in cell death and promote disease on the organismal level. Canonical DSB repair pathways, NHEJ and HR, have evolved to combat these lesions and restore genome integrity. When considering the types and extent of damage that can be associated with DSBs, it stands to reason that some breaks will not be accessible to either canonical repair pathway for a variety of reasons. Associated base damage (e.g. abasic sites or occluding protein adducts), or the nature of the break (e.g. “one-ended” breaks from replication fork collapse) may prevent loading of or repair by the NHEJ machinery. Similarly, cell cycle restrictions or failures during HR may result in free, resected DSB end structures that cannot be metabolized by NHEJ. For such contexts, cells must rely on an alternative means to repair damage.

My work, presented in this dissertation, shows that DNA Polymerase θ defines a third DSB repair pathway, termed TMEJ, that has an important role in protecting genome stability. In Chapter 2, I define the biological substrates and cellular mechanism of TMEJ by means of an extrachromosomal end joining assay and characterization of repair junctions. I then present in Chapter 3, data defining the role of TMEJ during chromosomal repair. I show here that TMEJ is essential for cell survival when canonical pathways are impaired or otherwise unable to resolve damage, and when DNA end resection is misregulated. This chapter also reveals a role for Pol θ in the mutagenic

signatures observed in hereditary breast cancer. Thus, my dissertation reveals TMEJ to be a flexible, efficient, semi-conservative DSB repair pathway that can have major impacts on organismal health.

4.1 Identification of biological substrates

Prior work has suggested that Pol θ promotes microhomology-mediated repair of DSBs in *D.melanogaster and C. elegans*, but has not directly identified the biological intermediates or mechanisms that lead to these resolutions[99,106,107]. My work, using an extrachromosomal end joining assay, shows that in mammalian cells, Pol θ directly utilizes 3' resected end structures as substrate for a DNA synthesis reaction and repair. Out of 17 mammalian DNA polymerases, this synthesis activity is unique to Pol θ. This finding is consistent with previously published work, where a purified polymerase domain fragment of Pol θ was capable of joining DNA substrates via a terminal microhomology alignment[135]. However, contrary to this *in vitro* work, where Pol θ activity was limited to synthesis of less than 15 bp, I have found that in cells, Pol θ efficiently mediates the repair of substrates requiring at least 70 nt of DNA synthesis. I also show that in cells, Pol θ does not dissociate at the ssDNA to duplex transition and is capable of bypassing 5' abasic site damage by displacing the top strand of an annealed duplex and continuing synthesis for at least an additional 15 nt. Taken together, these results show that in cells, Pol θ mediates a "long-patch" DSB repair process that is far more processive and efficient than has been previously described.

It is important to note however, that TMEJ is not the only option for joining resected substrates. When the extent of resection is 45 nt or less, nucleolytic removal of the 3'

overhang and joining by NHEJ is the dominant fate. The nuclease responsible for trimming these long 3' tails is currently unknown, but the clipping reaction appears to be tied to the NHEJ complex, with Artemis being a strong nuclease candidate. As resection extends past 45 nt, genetic requirements switch and end joining predominantly relies on Pol θ. This is explained by the inability of Ku to load on longer 3' ssDNA tails, perhaps as a result of intramolecular secondary structures formed on the long overhangs.

Extending the observation that impaired Ku loading effectively shunts substrate to TMEJ, I also show that Pol θ has a unique role in the joining of forked substrates that are mimetic of the one-ended DSB structures expected following replication fork collapse. These structures completely ablate the ability of Ku to load onto the broken end and therefore cannot be repaired by NHEJ. The unique ability of Pol θ to resolve these structures provides evidence that TMEJ is a crucial pathway in high replication stress environments. This hypothesis is corroborated by the observations that Pol θ is most highly expressed in rapidly dividing germ cells, and is essential for early embryonic development and germ cell maintenance in some model organisms[88,102,136].

4.2 Microhomology search mechanism

DNA end resection initiates with an MRN-dependent generation of ~20-50 nt of ssDNA and can extend to several kilobases in length[45,46,128]. End resection exposes patches of microhomology between the two DNA strands that are used to align strands and prime synthesis by Pol θ; yet it was not known how much complementarity is needed to prime synthesis or how microhomologies are found during cellular repair. Using high throughput sequencing of repair junctions from both extrachromosomal

substrate and CRISPR/Cas9 derived chromosomal DSBs, I show here that Pol θ-mediated joins are enriched over chance only when there is an alignment of greater than 2 bp. Chromosomally, Pol θ will use microhomologies as large as 7 nucleotides, though the chance of such tandem repeats of homology decreases exponentially with size, making the practical definition of microhomology in the 3-6 bp range.

My work also defines a role for the N-terminal helicase-like domain of Pol θ in an active search for microhomology alignments. Cells harboring an inactivating mutation in the Walker-A motif of the ATPase domain (K121M) have a reduced ability to utilize internally embedded microhomologies, and instead strongly favor alignment by a microhomology at the absolute 3' termini. This result is consistent with recent structural studies of the helicase-like domain which proposed a model where 3' DNA tails are pulled through a central channel and slid past each other[90]. This structure also revealed several new motifs, including an "RAR" motif that might directly interact with the 3' hydroxyl of the resected end. In light of these results, a full mechanistic exploration of the helicase-like domain and microhomology search engine should be undertaken.

Both extrachromosomal and chromosomal repair junctions recovered from wild type cells show that this microhomology search is restricted to the terminal ~25 nt of each 3' end. Employing an active search mechanism for alignments > 2 bp within this 2x25 nt space allows for TMEJ to repair virtually any pair of DSB ends, regardless of sequence context, in manner that limits deletion size to no more than 50 bp. Thus, the TMEJ mechanism protects the genome from the toxic effects of DSBs and limits catastrophically large deletion, though in a manner that is likely to be substantially more mutagenic than NHEJ or HR.

4.3 Other TMEJ Factors

Pol θ is an essential core factor for TMEJ/Alt-EJ/MMEJ, yet it is likely only one of a host of associated factors. My work indicates that Mre11 is also essential for TMEJ, with cells depleted of Mre11 via CRISPR/Cas9 mutagenesis acting as Pol θ knockouts during chromosomal end joining. This result is likely through an upstream role in end resection and the generation of Pol θ substrate. While these two factors begin to delineate the steps of the pathway, there is otherwise little information known about associated factors. Previous studies have suggested that both Ligase I and III participate in Alt-EJ[137]. This is consistent with my observations that repair by TMEJ involves strand displacement synthesis; ultimately this synthesis activity will result in a nick structure that must sealed by either Ligase I or III. The following questions raised by my work should be explored in further detail in an effort to identify other components of the pathway: what nucleases are associated with TMEJ, and does Pol θ associate with processivity factors during cellular repair?

My results show that Pol θ effectively uses embedded microhomology alignments that will generate 3' nonhomologous tails once the two ssDNA strands are aligned; however Pol θ is incapable of synthesizing from a mispaired nonhomologous tail, with efficiency reduced over 1000-fold on such a substrate. This indicates that a nuclease must resolve at least one of the tails prior to synthesis. For TMEJ to act efficiently, this nuclease activity must be tightly coupled with Pol θ. Candidates for this nuclease include Mre11 and XPF:ERCC1. In particular, the XPF ortholog Rad10 has previously described roles MMEJ in *S. cerevisiae*, while in a mammalian model, loss of ERCC1 hypersensitized Ku80-deficient cells to IR exposure[138,139]. It is also of interest to identify

the nuclease responsible for resolving the 5' flap generated by strand displacement synthesis. FEN-1 is a strong candidate for this, in line with its known role in trimming 5' flaps during Okazaki fragment resolution[140].

With respect to processivity factors, I have shown that Pol θ mediates the repair of much larger tracts of DNA than was previously suggested by *in vitro* studies[135]. It is possible that Pol θ associates with a factor such as PCNA or Rev1 during cellular repair, though there is no obvious PIP motif in the primary sequence of the protein. Another possibility is that Pol θ acts in a manner similar to translesion synthesis polymerases. In such a model, Pol θ would synthesize enough from a microhomology alignment that a stable primer was made for a replicative polymerase to switch on and complete repair.

4.4 Relationships between TMEJ and canonical double strand break repair pathways

Microhomology-mediated end joining was initially discovered in systems that were deficient in NHEJ[11-15,76]. These studies, coupled with the relatively mild radiosensitivity of Pol θ-deficient cells [107], point to a role for TMEJ as a backup repair pathway. I have extended these observations to show that under normal cellular conditions, Pol θ acts on approximately 5% of Cas9-induced double strand breaks; by comparison > 75% of such breaks are engaged by NHEJ. However, when cells are made deficient in NHEJ, TMEJ accounts for the majority of DSB repair. Further, cells that are doubly deficient in Ku70 and Pol θ have severe proliferation defects, including a reduced fraction of cells that actively replicate DNA and increased spontaneous chromosomal fusions. Taken

together, my results provide evidence that TMEJ is a minor, yet essential back up repair pathway to NHEJ.

In contrast, the relationship between HR and TMEJ is not well defined. The two pathways both utilize the resected end intermediate as substrate, yet it is not clear if there is an active competition for resected ends. My results indicate that TMEJ does not impact levels of cellular HR, when using a donor construct that has 500 bp of homology flanking the left and right sides of the DSB (for a total of 1 kb of homology). However, continuing experiments in our lab suggest there may be a mild competition for substrate when length of homologous donor, and thus extent of resection, is decreased to 250 bp on each side (for a total of 500 bp of homology). This is consistent with reports of a mild stimulation of HR in Pol θ-deficient cells when measuring recombination of a relatively small (500 bp) GFP reporter[89]. It is plausible that when end resection is shorter-range or slow, Pol θ may have preferential access to ends. If true, this might be explained by a reduced access of RPA – a factor that promotes repair by HR, yet suppresses MMEJ[141] – to more minimally resected ends. Further exploring the kinetics of resection will be essential to understanding the interplay between these two pathways.

Additional studies are also needed to assess whether loss of Pol θ is synthetic lethal with loss of HR factors in the same manner that it is with NHEJ. My work suggests that the mechanistic basis for synthetic lethality upon combined deficiency in TMEJ and NHEJ is that an excess of resected DSBs resulting from the loss of Ku70 or 53BP1 fail to be resolved. Deficiency in 53BP1 in particular leads to resection in G1 phase of the cell cycle, where HR is not able to act[61]. This hypothesis is supported by an increase in Rad51 foci in *Polq*$^{-/-}$ cells, presented here and by others[89,108]. Combined with the

observation that loss of Pol θ does not impact levels of successful HR, the increased number of foci suggests that these are free resected ends, bound by recombinase, but not productively involved in repair. In a similar manner, loss of later-stage HR factors such as BRCA2, Rad54, or Holliday junction resolvases may lead to an accumulation of resected ends that is also toxic in Pol θ-deficient cells.

4.5 Polymerase θ and genome stability

It is clear that under normal cellular conditions, the properly regulated use of TMEJ as a backup repair pathway protects genome integrity and cell viability, though the repair mechanism is inherently mutagenic. TMEJ limits potentially catastrophic deletion and provides an outlet for repairing potentially toxic double strand breaks. Pol θ also plays a role in protecting against major chromosomal rearrangement, though its role in protecting against translocations is only revealed when cells are also deficient in NHEJ. I show here that compound loss of Pol θ in *Ku70*$^{-/-}$ cells results in a 4-fold increase in translocations driven by tandem Cas9-induced DSBs on different chromosomes.

There are however, aspects of TMEJ that have the potential to be deleterious even under normal conditions. A rare class of TMEJ products that occur during chromosomal break repair involve large insertions at the repair junction that are templated from other chromosomes. Pol θ appears to jump between chromosomes, priming successive segments of synthesis from microhomologies. These pseudo-translocation products are structurally similar to microhomology-mediated break induced replication (MM-BIR) events that occur in *S. cerevisiae*. While these products are generally rare (1.5×10^{-3} in *Ku70*$^{-/-}$ cells), I was only able to quantify those products that returned to the site of the

original DSB and were thus amplified and sequenced. If a larger fraction of products jump chromosomes and do not return, it may represent a potent mechanism for loss of heterozygosity.

4.6 Polymerase θ as a therapeutic target

One of the central findings of my research is that Pol θ is a strong candidate for therapy in the treatment of hereditary breast cancers; especially those tumors which are resistant to PARPi therapy. This is supported by data showing that the chromosomal mutation signature of Pol θ (microhomology >2 bp, centered at long deletions or insertions) is identical to the somatic mutation signature seen in BRCA1/2-deficient hereditary breast cancers, (COSMIC signature 3[132]). This is consistent with previous studies indicating that Pol θ overexpression is associated with poor prognosis breast cancer, and that Pol θ suppression may reduce tumor volume and increase survival of HR-deficient tumors[89]. Strengthening this argument, I also show here that loss of Pol θ leads to synthetic lethality with loss of 53BP1. This anti-resection factor rescues growth and viability of BRCA1-deficient hereditary breast cancers, and protects the tumors from PARPi therapy[59,129]. My results suggest that Pol θ will be essential in this mode of therapy resistance and will be a potent target for the treatment of such cancers. Development of small molecule drugs either targeting Pol θ directly, or exploiting Pol θ activity has already begun. A recent study presents that size-expanded benzo-nucleotide analogs are incorporated exclusively by Pol θ and that they act as DNA chain-terminators upon incorporation[135].

4.7 Concluding Remarks

The work discussed within this thesis expands the current understanding of mammalian DSB repair, with direct consequences for the advancement of human health. My work and that of others defines the biological substrates, mechanism, context, and therapeutic potential of a novel DSB repair pathway mediated by DNA Polymerase θ. This work also provides a roadmap for the mutational processes associated with DSB repair and contributes to the advancement of the burgeoning field of directed DNA editing technologies. As our knowledge of Pol θ-directed DNA repair continues to expand it is becoming evident that TMEJ has broad implications beyond those of a simple back up repair pathway.

REFERENCES

1. Blunt, T. *et al.* Defective DNA-dependent protein kinase activity is linked to V(D)J recombination and DNA repair defects associated with the murine scid mutation. *Cell* **80,** 813–823 (1995).

2. Frank, K. M. *et al.* DNA Ligase IV Deficiency in Mice Leads to Defective Neurogenesis and Embryonic Lethality via the p53 Pathway. *Mol. Cell* **5,** 993–1002 (2000).

3. Grawunder, U., Zimmer, D., Fugmann, S., Schwarz, K. & Lieber, M. R. 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 (1998).

4. Gu, Y., Jin, S., Gao, Y., Weaver, D. T. & Alt, F. W. 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. U.S.A.* **94,** 8076–8081 (1997).

5. Buck, D. *et al.* Cernunnos, a novel nonhomologous end joining factor, is mutated in human immunodeficiency with microcephaly. *Cell* **124,** 287–299 (2006).

6. Helleday, T., Lo, J., van Gent, D. C. & Engelward, B. P. DNA double-strand break repair: from mechanistic understanding to cancer treatment. *DNA Repair (Amst.)* **6,** 923–935 (2007).

7. Valerie, K. & Povirk, L. F. Regulation and mechanisms of mammalian double-strand break repair. *Oncogene* **22,** 5792–5812 (2003).

8. Rothkamm, K., Krüger, I., Thompson, L. H. & Löbrich, M. Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol. Cell. Biol.* **23,** 5706–5715 (2003).

9. Heyer, W.-D., Ehmsen, K. T. & Liu, J. Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.* **44,** 113–139 (2010).

10. Chapman, J. R., Taylor, M. R. G. & Boulton, S. J. Playing the end game: DNA double-strand break repair pathway choice. *Mol. Cell* **47,** 497–510 (2012).

11. Perrault, R., Wang, H., Wang, M., Rosidi, B. & Iliakis, G. Backup pathways of NHEJ are suppressed by DNA-PK. *J. Cell. Biochem.* **92,** 781–794 (2004).

12. Wang, H. *et al.* Biochemical evidence for Ku-independent backup pathways of NHEJ. *Nucleic Acids Res.* **31,** 5377–5388 (2003).

13. Guirouilh-Barbat, J. *et al.* Impact of the KU80 pathway on NHEJ-induced genome rearrangements in mammalian cells. *Mol. Cell* **14,** 611–623 (2004).

14. Corneo, B. *et al.* Rag mutations reveal robust alternative end joining. *Nature* **449,** 483–486 (2007).

15. Kabotyanski, E. B., Gomelsky, L., Han, J. O., Stamato, T. D. & Roth, D. B. Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Res.* **26,** 5333–5342 (1998).

16. Hefferin, M. L. & Tomkinson, A. E. Mechanism of DNA double-strand break repair by non-homologous end joining. *DNA Repair (Amst.)* **4,** 639–648 (2005).

17. Blier, P. R., Griffith, A. J., Craft, J. & Hardin, J. A. Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. *J. Biol. Chem.* **268,** 7594–7601 (1993).

18. Falzon, M., Fewell, J. W. & Kuff, E. L. EBP-80, a transcription factor closely resembling the human autoantigen Ku, recognizes single- to double-strand transitions in DNA. *J. Biol. Chem.* **268,** 10546–10552 (1993).

19. Mimori, T., Hardin, J. A. & Steitz, J. A. Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. *J. Biol. Chem.* **261,** 2274–2278 (1986).

20. Mimori, T. & Hardin, J. A. Mechanism of interaction between Ku protein and DNA. *J. Biol. Chem.* **261,** 10375–10379 (1986).

21. Walker, J. R., Corpina, R. A. & Goldberg, J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412,** 607–614 (2001).

22. Roberts, S. A. & Ramsden, D. A. Loading of the nonhomologous end joining factor, Ku, on protein-occluded DNA ends. *J. Biol. Chem.* **282,** 10605–10613 (2007).

23. Roberts, S. A. *et al.* Ku is a 5'-dRP/AP lyase that excises nucleotide damage near broken ends. *Nature* **464,** 1214–1217 (2010).

24. Strande, N., Roberts, S. A., Oh, S., Hendrickson, E. A. & Ramsden, D. A. Specificity of the dRP/AP lyase of Ku promotes nonhomologous end joining (NHEJ) fidelity at damaged ends. *J. Biol. Chem.* **287,** 13686–13693 (2012).

25. Strande, N. T. *et al.* Requirements for 5'dRP/AP lyase activity in Ku. *Nucleic Acids Res.* **42,** 11136–11143 (2014).

26. DeFazio, L. G., Stansel, R. M., Griffith, J. D. & Chu, G. Synapsis of DNA ends by DNA-dependent protein kinase. *EMBO J.* **21,** 3192–3200 (2002).

27. Ding, Q. *et al.* Autophosphorylation of the catalytic subunit of the DNA-dependent protein kinase is required for efficient end processing during DNA double-strand break repair. *Mol. Cell. Biol.* **23,** 5836–5848 (2003).

28. Weterings, E., Verkaik, N. S., Brüggenwirth, H. T., Hoeijmakers, J. H. J. & van Gent, D. C. The role of DNA dependent protein kinase in synapsis of DNA ends. *Nucleic Acids Res.* **31,** 7238–7246 (2003).

29. Calsou, P., Delteil, C., Frit, P., Drouet, J. & Salles, B. Coordinated assembly of Ku and p460 subunits of the DNA-dependent protein kinase on DNA ends is necessary for XRCC4-ligase IV recruitment. *J. Mol. Biol.* **326,** 93–103 (2003).

30. Chen, L., Trujillo, K., Sung, P. & Tomkinson, A. E. Interactions of the DNA ligase IV-XRCC4 complex with DNA ends and the DNA-dependent protein kinase. *J. Biol. Chem.* **275,** 26196–26205 (2000).

31. Grawunder, U. *et al.* Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* **388,** 492–495 (1997).

32. Nick McElhinny, S. A., Snowden, C. M., McCarville, J. & Ramsden, D. A. Ku recruits the XRCC4-ligase IV complex to DNA ends. *Mol. Cell. Biol.* **20,** 2996–3003 (2000).

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

34. Wu, P.-Y. *et al.* Interplay between Cernunnos-XLF and nonhomologous end joining proteins at DNA ends in the cell. *J. Biol. Chem.* **282,** 31937–31943 (2007).

35. Waters, C. A. *et al.* The fidelity of the ligation step determines how ends are resolved during nonhomologous end joining. *Nat Commun* **5,** 4286 (2014).

36. Wang, J. C. Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol.* **3,** 430–440 (2002).

37. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362,** 709–715 (1993).

38. Demple, B. & Harrison, L. Repair of oxidative damage to DNA: enzymology and biology. *Annu. Rev. Biochem.* **63,** 915–948 (1994).

39. Ma, Y., Pannicke, U., Schwarz, K. & Lieber, M. R. 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 (2002).

40. Moshous, D. *et al.* 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 (2001).

41. Schellenberg, M. J. *et al.* Reversal of DNA damage induced Topoisomerase 2 DNA-protein crosslinks by Tdp2. *Nucleic Acids Res.* **44,** 3829–3844 (2016).

42. Nick McElhinny, S. A. *et al.* A gradient of template dependence defines distinct biological roles for family X polymerases in nonhomologous end joining. *Mol. Cell* **19,** 357–366 (2005).

43. DeRose, E. F. *et al.* Solution structure of polymerase mu's BRCT Domain reveals an element essential for its role in nonhomologous end joining. *Biochemistry* **46,** 12100–12110 (2007).

44. Pryor, J. M. *et al.* Essential role for polymerase specialization in cellular nonhomologous end joining. *Proc. Natl. Acad. Sci. U.S.A.* **112,** E4537–45 (2015).

45. Symington, L. S. End resection at double-strand breaks: mechanism and regulation. *Cold Spring Harb Perspect Biol* **6,** a016436–a016436 (2014).

46. Symington, L. S. Mechanism and regulation of DNA end resection in eukaryotes. *Crit. Rev. Biochem. Mol. Biol.* **51,** 195–212 (2016).

47. Shibata, A. *et al.* Factors determining DNA double-strand break repair pathway choice in G2 phase. *EMBO J.* **30,** 1079–1092 (2011).

48. Liao, S., Guay, C., Toczylowski, T. & Yan, H. Analysis of MRE11's function in the 5“-->3” processing of DNA double-strand breaks. *Nucleic Acids Res.* **40,** 4496–4506 (2012).

49. Shibata, A. *et al.* DNA double-strand break repair pathway choice is directed by distinct MRE11 nuclease activities. *Mol. Cell* **53,** 7–18 (2014).

50. Paull, T. T. & Gellert, M. Nbs1 potentiates ATP-driven DNA unwinding and endonuclease cleavage by the Mre11/Rad50 complex. *Genes Dev.* **13,** 1276–1288 (1999).

51. Paull, T. T. & Gellert, M. The 3‘ to 5’ exonuclease activity of Mre 11 facilitates repair of DNA double-strand breaks. *Mol. Cell* **1,** 969–979 (1998).

52. Moreau, S., Ferguson, J. R. & Symington, L. S. The nuclease activity of Mre11 is required for meiosis but not for mating type switching, end joining, or telomere maintenance. *Mol. Cell. Biol.* **19,** 556–566 (1999).

53. Limbo, O. *et al.* Ctp1 is a cell-cycle-regulated protein that functions with Mre11 complex to control double-strand break repair by homologous recombination. *Mol. Cell* **28,** 134–146 (2007).

54. Sartori, A. A. *et al.* Human CtIP promotes DNA end resection. *Nature* **450,** 509–514 (2007).

55. Anand, R., Ranjha, L., Cannavo, E. & Cejka, P. Phosphorylated CtIP Functions as a Co-factor of the MRE11-RAD50-NBS1 Endonuclease in DNA End Resection. *Mol. Cell* **64,** 940–950 (2016).

56. Aparicio, T. & Gautier, J. BRCA1-CtIP interaction in the repair of DNA double-strand breaks. *Mol Cell Oncol* **3,** e1169343 (2016).

57. Cruz-García, A., López-Saavedra, A. & Huertas, P. BRCA1 accelerates CtIP-mediated DNA-end resection. *Cell Rep* **9,** 451–459 (2014).

58. Nimonkar, A. V. *et al.* 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 (2011).

59. Bunting, S. F. *et al.* 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* **141,** 243–254 (2010).

60. Zimmermann, M., Lottersberger, F., Buonomo, S. B., Sfeir, A. & de Lange, T. 53BP1 regulates DSB repair using Rif1 to control 5' end resection. *Science* **339,** 700–704 (2013).

61. Bakr, A. *et al.* Impaired 53BP1/RIF1 DSB mediated end-protection stimulates CtIP-dependent end resection and switches the repair to PARP1-dependent end joining in G1. *Oncotarget* **7,** 57679–57693 (2016).

62. Escribano-Díaz, C. *et al.* A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. *Mol. Cell* **49,** 872–883 (2013).

63. Escribano-Díaz, C. & Durocher, D. DNA repair pathway choice--a PTIP of the hat to 53BP1. *EMBO Rep.* **14,** 665–666 (2013).

64. Langerak, P., Mejia-Ramirez, E., Limbo, O. & Russell, P. 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 (2011).

65. Fradet-Turcotte, A. *et al.* 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark. *Nature* **499,** 50–54 (2013).

66. Jacquet, K. *et al.* The TIP60 Complex Regulates Bivalent Chromatin Recognition by 53BP1 through Direct H4K20me Binding and H2AK15 Acetylation. *Mol. Cell* **62,** 409–421 (2016).

67. Aparicio, T., Baer, R. & Gautier, J. DNA double-strand break repair pathway choice and cancer. *DNA Repair (Amst.)* **19,** 169–175 (2014).

68. Chen, H., Lisby, M. & Symington, L. S. RPA coordinates DNA end resection and prevents formation of DNA hairpins. *Mol. Cell* **50,** 589–600 (2013).

69. Shivji, M. K. K. *et al.* A region of human BRCA2 containing multiple BRC repeats promotes RAD51-mediated strand exchange. *Nucleic Acids Res.* **34,** 4000–4011 (2006).

70. Carreira, A. *et al.* The BRC repeats of BRCA2 modulate the DNA-binding selectivity of RAD51. *Cell* **136,** 1032–1043 (2009).

71. Yu, D. S. *et al.* Dynamic control of Rad51 recombinase by self-association and interaction with BRCA2. *Mol. Cell* **12,** 1029–1041 (2003).

72. Esashi, F., Galkin, V. E., Yu, X., Egelman, E. H. & West, S. C. Stabilization of RAD51 nucleoprotein filaments by the C-terminal region of BRCA2. *Nat. Struct. Mol. Biol.* **14,** 468–474 (2007).

73. Davies, A. A. *et al.* Role of BRCA2 in control of the RAD51 recombination and DNA repair protein. *Mol. Cell* **7,** 273–282 (2001).

74. Zhao, W. *et al.* Promotion of BRCA2-Dependent Homologous Recombination by DSS1 via RPA Targeting and DNA Mimicry. *Mol. Cell* **59,** 176–187 (2015).

75. Renkawitz, J., Lademann, C. A., Kalocsay, M. & Jentsch, S. Monitoring homology search during DNA double-strand break repair in vivo. *Mol. Cell* **50,** 261–272 (2013).

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

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

78. Truong, L. N. *et al.* Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 7720–7725 (2013).

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

80. Lee, K. & Lee, S. E. Saccharomyces cerevisiae Sae2- and Tel1-dependent single-strand DNA formation at DNA break promotes microhomology-mediated end joining. *Genetics* **176,** 2003–2014 (2007).

81. Yan, C. T. *et al.* IgH class switching and translocations use a robust non-classical end joining pathway. *Nature* **449,** 478–482 (2007).

82. Lee-Theilen, M., Matthews, A. J., Kelly, D., Zheng, S. & Chaudhuri, J. CtIP promotes microhomology-mediated alternative end joining during class-switch recombination. *Nat. Struct. Mol. Biol.* **18,** 75–79 (2011).

83. Simsek, D. *et al.* DNA ligase III promotes alternative nonhomologous end joining during chromosomal translocation formation. *PLoS Genet.* **7,** e1002080 (2011).

84. Della-Maria, J. *et al.* Human Mre11/human Rad50/Nbs1 and DNA ligase IIIalpha/XRCC1 protein complexes act together in an alternative nonhomologous end joining pathway. *J. Biol. Chem.* **286,** 33845–33853 (2011).

85. Wang, M. *et al.* PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res.* **34,** 6170–6182 (2006).

86. Audebert, M., Salles, B. & Calsou, P. 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 (2004).

87. Yousefzadeh, M. J. & Wood, R. D. DNA polymerase POLQ and cellular defense against DNA damage. *DNA Repair (Amst.)* **12,** 1–9 (2013).

88. Seki, M., Marini, F. & Wood, R. D. POLQ (Pol theta), a DNA polymerase and DNA-dependent ATPase in human cells. *Nucleic Acids Res.* **31,** 6117–6126 (2003).

89. Ceccaldi, R. *et al.* Homologous-recombination-deficient tumours are dependent on Polθ-mediated repair. *Nature* **518,** 258–262 (2015).

90. Newman, J. A., Cooper, C. D. O., Aitkenhead, H. & Gileadi, O. Structure of the Helicase Domain of DNA Polymerase Theta Reveals a Possible Role in the Microhomology-Mediated End joining Pathway. *Structure* **23,** 2319–2330 (2015).

91. Beagan, K. & McVey, M. Linking DNA polymerase theta structure and function in health and disease. *Cell. Mol. Life Sci.* **73,** 603–615 (2016).

92. Zahn, K. E., Averill, A. M., Aller, P., Wood, R. D. & Doublié, S. Human DNA polymerase θ grasps the primer terminus to mediate DNA repair. *Nat. Struct. Mol. Biol.* **22,** 304–311 (2015).

93. Arana, M. E., Seki, M., Wood, R. D., Rogozin, I. B. & Kunkel, T. A. Low-fidelity DNA synthesis by human DNA polymerase theta. *Nucleic Acids Res.* **36,** 3847–3856 (2008).

94. Hogg, M., Seki, M., Wood, R. D., Doublié, S. & Wallace, S. S. Lesion bypass activity of DNA polymerase θ (POLQ) is an intrinsic property of the pol domain and depends on unique sequence inserts. *J. Mol. Biol.* **405,** 642–652 (2011).

95. Seki, M. *et al.* High-efficiency bypass of DNA damage by human DNA polymerase Q. *EMBO J.* **23,** 4484–4494 (2004).

96. Hogg, M., Sauer-Eriksson, A. E. & Johansson, E. Promiscuous DNA synthesis by human DNA polymerase θ. *Nucleic Acids Res.* **40,** 2611–2622 (2012).

97. Shima, N. *et al.* Phenotype-based identification of mouse chromosome instability mutants. *Genetics* **163,** 1031–1040 (2003).

98. Shima, N., Munroe, R. J. & Schimenti, J. C. The mouse genomic instability mutation chaos1 is an allele of Polq that exhibits genetic interaction with Atm. *Mol. Cell. Biol.* **24,** 10381–10389 (2004).

99. Chan, S. H., Yu, A. M. & McVey, M. Dual roles for DNA polymerase theta in alternative end joining repair of double-strand breaks in Drosophila. *PLoS Genet.* **6,** e1001005 (2010).

100. Yu, A. M. & McVey, M. Synthesis-dependent microhomology-mediated end joining accounts for multiple types of repair junctions. *Nucleic Acids Res.* **38,** 5706–5717 (2010).

101. Boyd, J. B., Sakaguchi, K. & Harris, P. V. mus308 mutants of Drosophila exhibit hypersensitivity to DNA cross-linking agents and are defective in a deoxyribonuclease. *Genetics* **125,** 813–819 (1990).

102. van Schendel, R., Roerink, S. F., Portegijs, V., van den Heuvel, S. & Tijsterman, M. Polymerase Θ is a key driver of genome evolution and of CRISPR/Cas9-mediated mutagenesis. *Nat Commun* **6,** 7394 (2015).

103. Roerink, S. F., van Schendel, R. & Tijsterman, M. Polymerase theta-mediated end joining of replication-associated DNA breaks in C. elegans. *Genome Res.* **24,** 954–962 (2014).

104. Koole, W. *et al.* A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. *Nat Commun* **5,** 3216 (2014).

105. Lemmens, B., van Schendel, R. & Tijsterman, M. Mutagenic consequences of a single G-quadruplex demonstrate mitotic inheritance of DNA replication fork barriers. *Nat Commun* **6,** 8909 (2015).

106. van Schendel, R., van Heteren, J., Welten, R. & Tijsterman, M. Genomic Scars Generated by Polymerase Theta Reveal the Versatile Mechanism of Alternative End joining. *PLoS Genet.* **12,** e1006368 (2016).

107. Yousefzadeh, M. J. *et al.* Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. *PLoS Genet.* **10,** e1004654 (2014).

108. Mateos-Gomez, P. A. *et al.* Mammalian polymerase θ promotes alternative NHEJ and suppresses recombination. *Nature* **518,** 254–257 (2015).

109. Wood, R. D. & Doublié, S. DNA polymerase θ (POLQ), double-strand break repair, and cancer. *DNA Repair (Amst.)* **44,** 22–32 (2016).

110. Lemée, F. *et al.* DNA polymerase theta up-regulation is associated with poor survival in breast cancer, perturbs DNA replication, and promotes genetic instability. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 13390–13395 (2010).

111. Brandalize, A. P. C. *et al.* A DNA repair variant in POLQ (c.-1060A > G) is associated to hereditary breast cancer patients: a case-control study. *BMC Cancer* **14,** 850 (2014).

112. Dai, C.-H. *et al.* Co-inhibition of pol θ and HR genes efficiently synergize with cisplatin to suppress cisplatin-resistant lung cancer cells survival. *Oncotarget* **7,** 65157–65170 (2016).

113. Kass, E. M. & Jasin, M. Collaboration and competition between DNA double-strand break repair pathways. *FEBS Lett.* **584,** 3703–3708 (2010).

114. Rassool, F. V. & Tomkinson, A. E. Targeting abnormal DNA double strand break repair in cancer. *Cell. Mol. Life Sci.* **67,** 3699–3710 (2010).

115. Thompson, L. H. Recognition, signaling, and repair of DNA double-strand breaks produced by ionizing radiation in mammalian cells: the molecular choreography. *Mutat. Res.* **751,** 158–246 (2012).

116. Decottignies, A. Alternative end joining mechanisms: a historical perspective. *Front Genet* **4,** 48 (2013).

117. Boboila, C., Alt, F. W. & Schwer, B. Classical and alternative end joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. *Adv. Immunol.* **116,** 1–49 (2012).

118. Yoon, J.-H., Roy Choudhury, J., Park, J., Prakash, S. & Prakash, L. A role for DNA polymerase θ in promoting replication through oxidative DNA lesion, thymine glycol, in human cells. *J. Biol. Chem.* **289,** 13177–13185 (2014).

119. Yoshimura, M. *et al.* Vertebrate POLQ and POLbeta cooperate in base excision repair of oxidative DNA damage. *Mol. Cell* **24,** 115–125 (2006).

120. Prasad, R. *et al.* Human DNA polymerase theta possesses 5'-dRP lyase activity and functions in single-nucleotide base excision repair in vitro. *Nucleic Acids Res.* **37,** 1868–1877 (2009).

121. Fernandez-Vidal, A. *et al.* A role for DNA polymerase θ in the timing of DNA replication. *Nat Commun* **5,** 4285 (2014).

122. Harris, P. V. *et al.* Molecular cloning of Drosophila mus308, a gene involved in DNA cross-link repair with homology to prokaryotic DNA polymerase I genes. *Mol. Cell. Biol.* **16,** 5764–5771 (1996).

123. Muzzini, D. M., Plevani, P., Boulton, S. J., Cassata, G. & Marini, F. Caenorhabditis elegans POLQ-1 and HEL-308 function in two distinct DNA interstrand cross-link repair pathways. *DNA Repair (Amst.)* **7,** 941–950 (2008).

124. Kent, T., Chandramouly, G., McDevitt, S. M., Ozdemir, A. Y. & Pomerantz, R. T. Mechanism of microhomology-mediated end joining promoted by human DNA polymerase θ. *Nat. Struct. Mol. Biol.* **22,** 230–237 (2015).

125. Maga, G., Shevelev, I., Ramadan, K., Spadari, S. & Hübscher, U. DNA polymerase theta purified from human cells is a high-fidelity enzyme. *J. Mol. Biol.* **319,** 359–369 (2002).

126. Sfeir, A. & Symington, L. S. Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? *Trends Biochem. Sci.* **40,** 701–714 (2015).

127. Higgins, G. S. *et al.* Overexpression of POLQ confers a poor prognosis in early breast cancer patients. *Oncotarget* **1,** 175–184 (2010).

128. Zhou, Y. & Paull, T. T. Direct measurement of single-stranded DNA intermediates in mammalian cells by quantitative polymerase chain reaction. *Anal. Biochem.* **479,** 48–50 (2015).

129. Bouwman, P. *et al.* 53BP1 loss rescues BRCA1 deficiency and is associated with triple-negative and BRCA-mutated breast cancers. *Nat. Struct. Mol. Biol.* **17,** 688–695 (2010).

130. Pierce, A. J., Hu, P., Han, M., Ellis, N. & Jasin, M. Ku DNA end-binding protein modulates homologous repair of double-strand breaks in mammalian cells. *Genes Dev.* **15,** 3237–3242 (2001).

131. Weinstock, D. M., Brunet, E. & Jasin, M. Formation of NHEJ-derived reciprocal chromosomal translocations does not require Ku70. *Nat. Cell Biol.* **9,** 978–981 (2007).

132. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500,** 415–421 (2013).

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

134. Sakofsky, C. J. *et al.* Translesion Polymerases Drive Microhomology-Mediated Break-Induced Replication Leading to Complex Chromosomal Rearrangements. *Mol. Cell* **60,** 860–872 (2015).

135. Kent, T. *et al.* DNA polymerase θ specializes in incorporating synthetic expanded-size (xDNA) nucleotides. *Nucleic Acids Res.* **44,** 9381–9392 (2016).

136. Thyme, S. B. & Schier, A. F. Polq-Mediated End Joining Is Essential for Surviving DNA Double-Strand Breaks during Early Zebrafish Development. *Cell Rep* **15,** 1611–1613 (2016).

137. Paul, K. *et al.* DNA ligases I and III cooperate in alternative non-homologous end joining in vertebrates. *PLoS ONE* **8,** e59505 (2013).

138. Ahmad, A. *et al.* ERCC1-XPF endonuclease facilitates DNA double-strand break repair. *Mol. Cell. Biol.* **28,** 5082–5092 (2008).

139. Haber, J. E. Transpositions and translocations induced by site-specific double-strand breaks in budding yeast. *DNA Repair (Amst.)* **5,** 998–1009 (2006).

140. Zheng, L. & Shen, B. Okazaki fragment maturation: nucleases take centre stage. *J Mol Cell Biol* **3,** 23–30 (2011).

141. Deng, S. K., Gibb, B., de Almeida, M. J., Greene, E. C. & Symington, L. S. RPA antagonizes microhomology-mediated repair of DNA double-strand breaks. *Nat. Struct. Mol. Biol.* **21,** 405–412 (2014).