5-2015

DNA POLYMERASE θ (POLQ) AND THE CELLULAR DEFENSE AGAINST DNA DAMAGE

Matthew J. Yousefzadeh

DNA POLYMERASE θ (POLQ) AND THE CELLULAR DEFENSE AGAINST DNA DAMAGE

BY

Matthew James Yousefzadeh, Ph.D. Candidate

APPROVED:

_____

Richard D. Wood, Ph.D., Supervisory Professor

_____

Rick A. Finch, Ph.D.

_____

Randy J. Legerski, Ph.D.

_____

Kevin M. McBride, Ph.D.

_____

Karen M. Vasquez, Ph.D.

APPROVED:

_____

Dean, The University of Texas
Graduate School of Biomedical Sciences at Houston

DNA POLYMERASE θ (POLQ) AND THE CELLULAR DEFENSE AGAINST DNA DAMAGE

A

Thesis

Presented to the Faculty of

The University of Texas Health Science Center at Houston

And

The University of Texas M.D. Anderson Cancer Center

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the degree of

DOCTOR OF PHILOSOPHY

By

Matthew James Yousefzadeh, Ph.D. Candidate

Houston, Texas

May 2015

DEDICATION

This thesis is dedicated to the following individuals:

My grandmother for teaching me that what you do matters more than what you say

My father for teaching me to dream

My mother for teaching me to care

My sister Emily for teaching me to never quit

"Service to others is the rent you pay for your room here on Earth."
Muhammad Ali

ACKNOWLEDGEMENTS

I would like to acknowledge the love and support of my family (Mom, Dad, Emily, Charlye, Uncle Steve, Aunt Diane, Aunt Boo, Brian, Little Bryan, Sam, Taylor, Justin, Kara, Liza, Nancy, and Henry J), my best friend Spencer Reed and his brother Brian Reed, my mentor Rick for providing an intellectually stimulating and safe place for a scientist to grow and all the great people who have passed through his lab during my time there (Karen, Kei-ichi, Sabine, Ella, Dave, Shelley, Junya, Maria, Mandira, Sarita, Megan, Leslie, Lidza, Becca, and Priyanka), my other committee members (Drs. Kevin McBride, Karen Vasquez, Rick Finch, Randy Legerski, Rodney Nairn, and Feng Wang-Johanning), Becky Brooks (for being a surrogate mother to myself and all the other graduate students), members of the McBride lab (Sean, Monika, Marc, Yun, and Josh), members of the Ramsden Lab (Dale, David, John, and Crystal), the fine folks at Molecular Biology Core (JJ, Yoko, Luis, Melissa, Angela, Madison, and Yuping) and Research Animal Support Facility (Dale, Lezlee, Donna, Barbara, Pam, Meredith, Debra, and Christina) for all the help throughout the years, Margie and Tilda (who I extend my utmost thanks and sincerest apologies for making loads of dirty dishes), Dr. Colin Brooks (who keeps Science Park from falling apart), Sheri Axtel, Avery Barksdale, Darci Long, all the wonderful folks in Technical Services (John, Bryan, Duke, Jen, and Chris) and Research Graphics (Chris and Joi), Patsy, Matt Morrin, Karl Zahn, Dr. Sylvie Doublié, Dr. Sankar Mitra, Arijit Dutta, Albino Bacolla, Wade Reh, Anirban Mukherjee, Dr. Mike Mitas (wherever the hell you are), Amanda Graham, Dr. Craig Blum, Sabri El-Naggar, “Little” Jeff Hall, “Big” Jeff Godfrey, “Medium” Jeff Tomlinson, Jim Deavor, Rick Heldrich, Andy Gelasco, Justin Wyatt, Dawne Taylor, Amy Rogers, Elizabeth Martin, Brian Spangler, Gamil Guirgis, Neil Tonks, Tom Cuff, Don Barrett, D. Reid Wiseman, Bob Frankis, Mark Lazarro, Peyre Pringle, Duncan Munro, Witold Witkowski, Nalini Mayroo, Richard Eik, Julie Pfaff, Robert Hopkins, Mr. Smiley, Coach Douty, Mrs. Wroble, Skip and Catherine Jakaitis, Alan and Candace Furness, Jim Hoy (for giving me my first real job) and Joe Campbell (for convincing

Jim not to fire me from my first job), Robert Ellington, the Assey family (Big Michael, Little Michael, Valerie, Amanda, Tim, and Stella), the Semsars (Mehdi, Mary, Ashley, Joe, John, and Lauren), Drs. Francesca Cole and Andrew Bergemann and their daughter Rohan, the Mitchells (David, Judy, and Josh), Drs. Mark Bedford and David Johnson (for being my unofficial mentors), Dr. John DiGiovanni (for that rotation that I never did), Michelle Bolner (my first friend at Science Park) and her even more awesome husband Bryan Frederick, Katie Reeh (my first Cuban friend) and her fantastic husband Mark, Aimee, Nicholas, Sitaram, Andria, Lan, Daric, Cari, Andy Sallinger, Garlen Yeager, Briana Dennehey (for her extensive editing), Joe Rodriguez, Mr. Tony, Nigel, Gary, Kevin, David, and in memory of Sue Jenkins Mauney, Rachel Mauney Yousefzadeh, Kelsey McGahan, Dr. Frank Kinard, Dr. Gary Asleson, and Dr. Ryan Monfeli (all people that were taken before their time and who made the world just that much of a better place).

DNA POLYMERASE θ (POLQ) AND THE CELLULAR DEFENSE AGAINST DNA DAMAGE

Matthew James Yousefzadeh, Ph.D. Candidate

Supervisory Professor: Richard D. Wood, Ph.D.

In mammalian cells, DNA polymerase θ (POLQ) is an unusual specialized DNA polymerase whose *in vivo* function is under active investigation. The protein is comprised of an N-terminal helicase-like domain, a C-terminal DNA polymerase domain, and a large central domain that spans between the two. This arrangement is also found in the *Drosophila* Mus308 protein, which helps confer resistance to DNA interstrand crosslinking agents. Homologs of POLQ and Mus308 are found in eukaryotes, including plants, but a comparison of phenotypes suggests that not all of these genes are functional orthologs. Flies with defective Mus308 are sensitive to DNA interstrand crosslinking agents, while mammalian cells with defective POLQ are primarily sensitive to DNA double-strand breaking agents. Cells from *Polq*-null mice are hypersensitive to radiation and the peripheral blood cells of these mice display increased spontaneous and ionizing radiation-induced levels of micronuclei (a hallmark of gross chromosomal aberrations), though the mice apparently develop normally.

Although a defect in the DNA polymerase POLQ leads to ionizing radiation sensitivity in mammalian cells, the relevant enzymatic pathway has not been identified. Here we define the specific mechanism by which POLQ restricts harmful DNA instability. Our experiments show that *Polq*-null murine cells are selectively hypersensitive to DNA strand-breaking agents, and that damage resistance requires the DNA polymerase activity of POLQ. Using a DNA break end joining assay in cells, the repair of DNA ends with long 3′ single-stranded overhangs was monitored. End joining events that retained much of the overhang were

dependent on POLQ, and independent of Ku70. To analyze this repair function in more detail, immunoglobulin class switch joining between DNA segments in antibody genes was examined. POLQ participates in the end joining of a DNA break during immunoglobulin class-switching, producing insertions of base pairs at the joins with homology to *IgH* switch-region sequences. Biochemical experiments with purified human POLQ protein revealed the mechanism generating the insertions during DNA end joining, relying on the unique ability of POLQ to extend DNA from minimally paired primers.

DNA breaks at the *IgH* locus can sometimes join with breaks in *Myc*, creating a chromosome translocation. A marked increase in *Myc/IgH* translocations was observed in *Polq*-defective mice, showing that POLQ suppresses genomic instability and genome rearrangements originating at DNA double-strand breaks. This work clearly defines a role and mechanism for mammalian POLQ in an alternative end joining pathway (termed synthesis-dependent end joining) that suppresses the formation of chromosomal translocations. Our findings depart from the prevailing view that alternative end joining processes are generically translocation-prone.

Class switch and junction analysis was also performed in mice lacking POLN, another DNA polymerase related to POLQ. I observed that POLN does not operate in the same alternative end joining pathway as does POLQ. Loss of *Poln* does not enhance the DNA damage hypersensitivity seen in cells lacking *Polq*. These findings suggest that while these two polymerases are structurally related they appear to have distinct functions in the cell. Analysis of the *Poln* phenotype is still ongoing. Further analysis of POLN and POLQ is required to clarify the mechanism by which they function in the cell.

TABLE OF CONTENTS

	PAGE
APPROVAL PAGE	**i**
TITLE PAGE	**ii**
DEDICATION	**iii**
ACKNOWLEDGEMENTS	**iv**
ABSTRACT	**vi**
TABLE OF CONTENTS	**vii**
LIST OF FIGURES	**xi**
LIST OF TABLES	**xiii**
LIST OF ABBREVIATIONS	**xiv**
CHAPTER 1: DNA POLYMERASE POLQ AND THE CELLULAR DEFENSE AGAINST DNA DAMAGE	**1**
1.1 Introduction	1
1.2 POLQ and DNA repair	17
1.3 Aims of this thesis	26
CHAPTER 2: MECHANISM OF SUPPRESSION OF CHROMOSOMAL INSTABILITY BY DNA POLYMERASE POLQ	**28**
2.1 Introduction	28
2.2 Material and methods	32
2.3 Results	45
2.4 Conclusions	75

CHAPTER 3: DOES A POLQ "SIGNATURE" EXIST IN TRANSLOCATIONS AND IMMUNOGLOBULIN V REGIONS **82**

3.1 Introduction 82

3.2 Material and methods 85

3.3 Results and future work 86

CHAPTER 4: THE EFFECT OF POLQ SUPPRESSION ON RADIOSENSITIVITY OF BREAST CANCER CELLS **89**

4.1 Introduction 89

4.2 Material and methods 91

4.3 Results and future work 94

CHAPTER 5: A-FAMILY DNA POLYMERASES AND THE REPAIR OF X-RAY INDUCED DNA DAMAGE **99**

5.1 Introduction 99

5.2 Material and methods 101

5.3 Results 105

5.4 Conclusions and future directions 114

CHAPTER 6: PREPARATION AND VALIDATION OF HUMAN POLQ ANTIBODIES **117**

6.1 Introduction 117

6.2 Material and methods 118

6.3 Results and discussion 124

6.4 Conclusion and future directions 136

CHAPTER 7: FUTURE DIRECTIONS **137**

7.1 POLQ maintains genomic integrity through its altEJ activity 137

7.2 POLQ diversifies the immunological portfolio 140

7.3 The impact of POLQ on breast cancer and chemotherapeutics 143

7.4 An “origin”al role for POLQ in DNA replication 146

7.5 Conclusion 147

Appendix 1. Analysis of REV7 in breast cancer and DNA damage sensitivity **148**

Appendix 2. Analysis of *Fam35a* expression in murine tissues **153**

BIBLIOGRAPHY **155**

VITA **182**

LIST OF FIGURES

Figure 1. Structural comparison and related grouping of *POLQ* gene family members 2

Figure 2. Map and alignment of POLQ polymerase domain 5

Figure 3. DNA double-strand break repair pathways 29

Figure 4. Immunoglobulin class-switch recombination 30

Figure 5. Hypersensitivity of *Polq*$^{-/-}$ bone marrow stromal cells to DNA strand-breaking agents 46

Figure 6. Loss of *Polq* contributes to chromosomal instability both spontaneously and in the presence of DNA damage 51

Figure 7. Complementation of the polymerase activity of POLQ rescues DNA damage hypersensitivity in cells lacking *Polq* 53

Figure 8. Complementation of polymerase-dead POLQ does not rescue DNA damage sensitivity 55

Figure 9. Analysis of DNA double-strand breaks and micronuclei in complemented *Polq* MEFs 56

Figure 10. Full-length *chaosI* mutant protein is poorly expressed 57

Figure 11. Insertions >1 nt at CSR junctions are *Polq*-dependent 60

Figure 12. End joining with extrachromosomal substrates 64

Figure 13. Unique template dependent DNA polymerase activity of POLQ 67

Figure 14. POLQ promiscuously extends single-stranded DNA ends through intermolecular templating 71

Figure 15. POLQ suppresses chromosomal translocation *in vivo* 73

Figure 16. Absolute quantification of transcript number shows that *Polq* is broadly expressed in tissues 75

Figure 17. Mechanism of insertion formation by POLQ during double-strand break repair 79

Figure 18. Survey of *POLQ* and DNA repair gene expression in breast cancer cells 95

Figure 19. shRNA suppression of POLQ in cancer cell lines 97

Figure 20. Absolute quantification of transcript number shows that *Poln* is highly expressed in tissues 105

Figure 21. Analysis of CSR junctions in *Poln* mice 107

Figure 22. Loss of *Poln* or *Polq* does not dramatically alter DNA repair capacity in mouse testis 110

Figure 23. Loss of *Poln* does not enhance DNA damage hypersensitivity found in *Polq* MEFs 111

Figure 24. Ectopic expression of POLQ in cancer cells induces DNA damage resistance 113

Figure 25. Screening of human POLQ antibodies 126

Figure 26. Epitope mapping of human POLQ antibodies 130

Figure 27. Screening of human POLQ rabbit polyclonal antibodies from Cell Signaling Technologies 132

Figure 28. Screening of human POLQ rabbit polyclonal antibodies from Bethyl Laboratories 134

Figure 29. Cell staining and immunofluorescence with human POLQ antibodies 135

Figure 30. Ectopic expression of REV7 can induce DNA damage hypersensitivity in breast cancer cells 149

Figure 31. Absolute quantification of transcript number shows that *Fam35a* is broadly expressed in tissues 153

LIST OF TABLES

Table 1. Conservation of POLQ “central” domain length and polymerase domain inserts in evolution 10

Table 2. Sensitivity of POLQ-defective cells to DNA damaging agents 19

Table 3. *POLQ* transcripts are found throughout the body 25

Table 4. Sequence composition of >1 nucleotide CSR insertions in *Polq* mice 62

Table 5. POLQ-dependent extension of 3’ DNA ends relies upon homology 70

Table 6. Sequencing of *IgH* and *IgK* sequences from *Polq* mice 87

Table 7. Sequence composition of >1 nucleotide CSR insertions in *Poln* mice 108

Table 8: Human POLQ antibodies 127

LIST OF ABBREVIATIONS

altEJ	alternative end joining
AP	abasic site
BER	base excision repair
BMSC	bone marrow stromal cell
cNHEJ	classical non-homologous end joining
CDR	complementarity determining region
CSR	class switch recombination
DNA	deoxyribonucleic acid
DSB	double-strand break
ENU	N-ethyl-N-nitrosourea
FACS	fluorescence-activated cell sorting
HR	homologous recombination
ICL	interstrand crosslink
Ig	immunoglobulin
Indel	nucleotide insertion and/or deletion
IR	ionizing radiation
MEF	mouse embryonic fibroblast
MMC	mitomycin c
MMEJ	microhomology-mediated end joining
MMS	methyl methane sulfonate
MN	micronuclei
Pol	DNA polymerase
SD-EJ	synthesis dependent end joining; also known as SD-MMEJ
Tg	thymine glycol
TLS	translesion DNA synthesis

TNBC	triple negative breast cancer
V region	variable region (of the immunoglobulin locus)
6-4 PP	6-4 photoproduct

Chapter 1: DNA polymerase POLQ and cellular defense against DNA damage

Note this chapter is based on: Yousefzadeh, M.J., and Wood, R.D. (2013). DNA polymerase POLQ and cellular defense against DNA damage. DNA Repair *12*, 1-9. With permission from the copyright holder.

1.1 Introduction

The functions of the 16 known mammalian DNA polymerases are currently being explored (Garcia-Gomez et al., 2013; Hubscher et al., 2002; Lange et al., 2011). DNA polymerases (pols) play pivotal roles not only in DNA replication (pols α, δ, ε), but also in base excision repair (pol β), mitochondrial replication and repair (pol γ), non-homologous end-joining and immunological diversity (pols λ, μ, and terminal-deoxynucleotidyl transferase), and DNA damage tolerance including translesion synthesis (η, κ, ζ, Rev1). Some DNA polymerases have roles in more than one pathway of DNA processing.

Identification of mammalian POLQ initially arose from interest in the Mus308 gene product of the fruit fly *Drosophila melanogaster*. *Mus308* mutants are hypersensitive to agents that cause DNA interstrand cross-links (ICL), with only modest sensitivity to monofunctional alkylating agents (Aguirrezabalaga et al., 1995; Harris et al., 1996). This suggested that Mus308 might play a specific role in repair of ICLs in DNA. Characterization of the *Mus308* gene showed that it encodes an unusual domain configuration, with the C-terminal part of the protein encoding a DNA polymerase, and the N-terminal part of the protein encoding a DNA helicase (Harris et al., 1996) (**Figure 1**).

A cDNA encoding the DNA polymerase domain of a homologous enzyme, designated DNA polymerase θ (POLQ), was identified by Sharief *et al.* (Sharief et al., 1999). The predicted protein of 1762 amino acids did not include a helicase-like domain, and is

now recognized to correspond to the C-terminal portion of full-length POLQ. Independently, Abbas and Linn assembled a larger 9.1 kb cDNA encoding both polymerase and helicase domains and predicted it to encode a protein of 2724 amino acids (Abbas, 2000). Seki *et al.* independently isolated the first complete and functional POLQ cDNA, showing that it encodes a 2590 amino acid protein with both a DNA polymerase and helicase-like domain (Seki et al., 2003). An orthologous mouse POLQ of 2544 amino acids was predicted from the genomic sequence by Shima *et al.* (Shima et al., 2003).

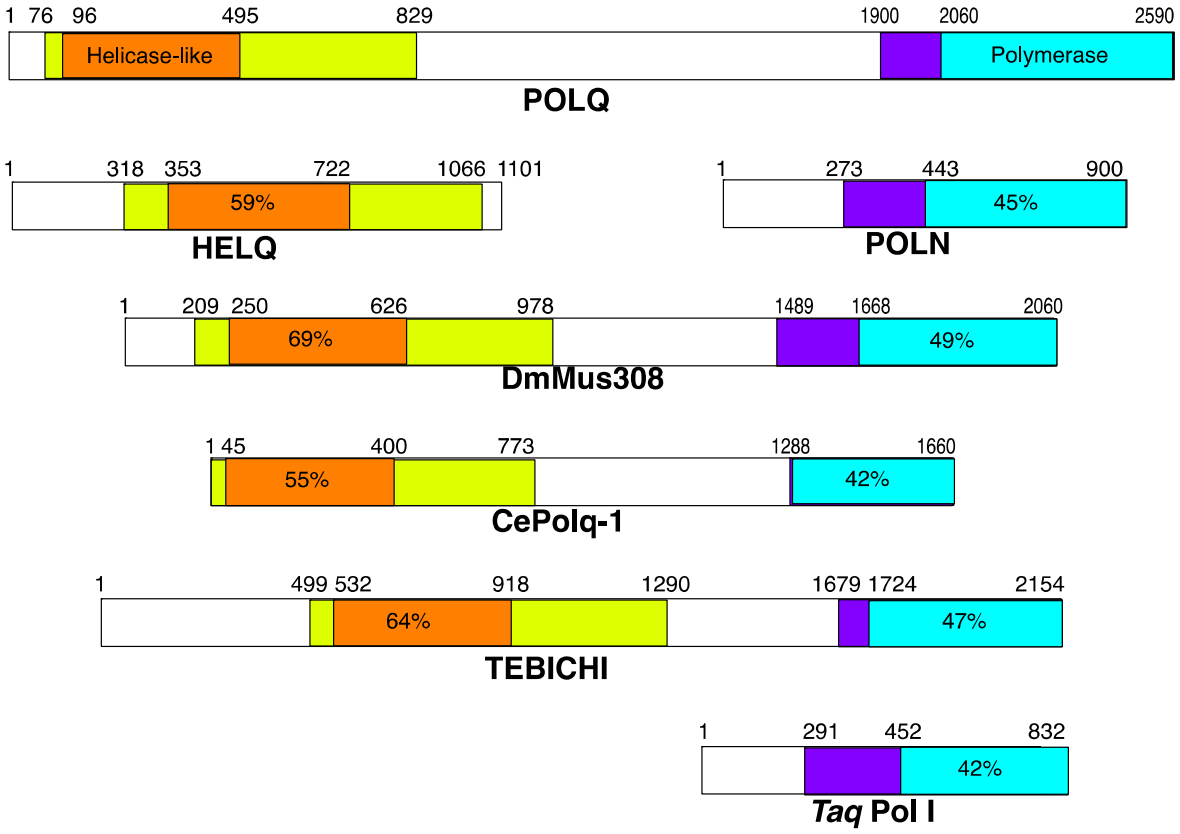


Figure 1. Structural comparison and related grouping of *POLQ* gene family members. The boundaries of the conserved domains and core catalytic domains of the helicase-like (yellow/orange) and polymerase (purple/blue) domains are noted in color and determined

through the Conserved Domain Database (Marchler-Bauer et al., 2011). Numbers above the proteins indicate amino acid positions. Percentages indicating the individual % similarity between the helicase-like and polymerase domains of POLQ were generated in MacVector.

The polymerase domains of POLQ and Mus308 belong to the "A" family of DNA polymerases, showing sequence similarity to *E. coli* DNA polymerase I (Harris et al., 1996; Marini et al., 2003). As described in more detail below, however, the polymerase domain of POLQ contains three unique "insert" regions within the DNA polymerase sequence (**Figure 2A and B**), which determine some of its biochemical properties (Seki et al., 2004). While much is known about the enzymatic properties of POLQ *in vitro*, further study is needed to better determine its functions in cells. Below I review data on the published literature regarding the function of POLQ and role in DNA repair.

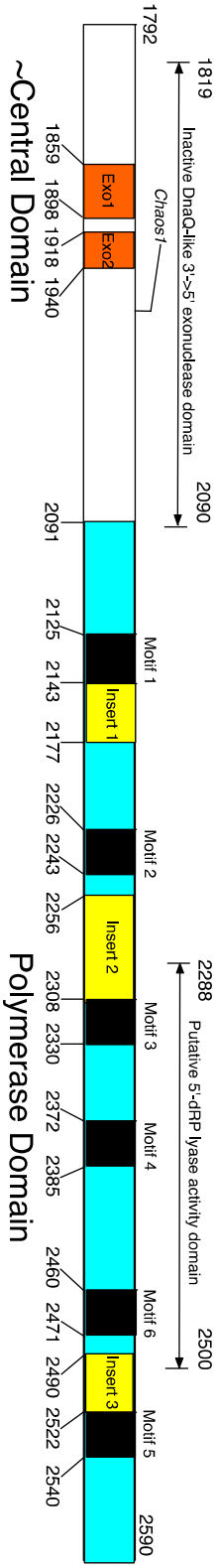
POLQ gene family distribution in nature

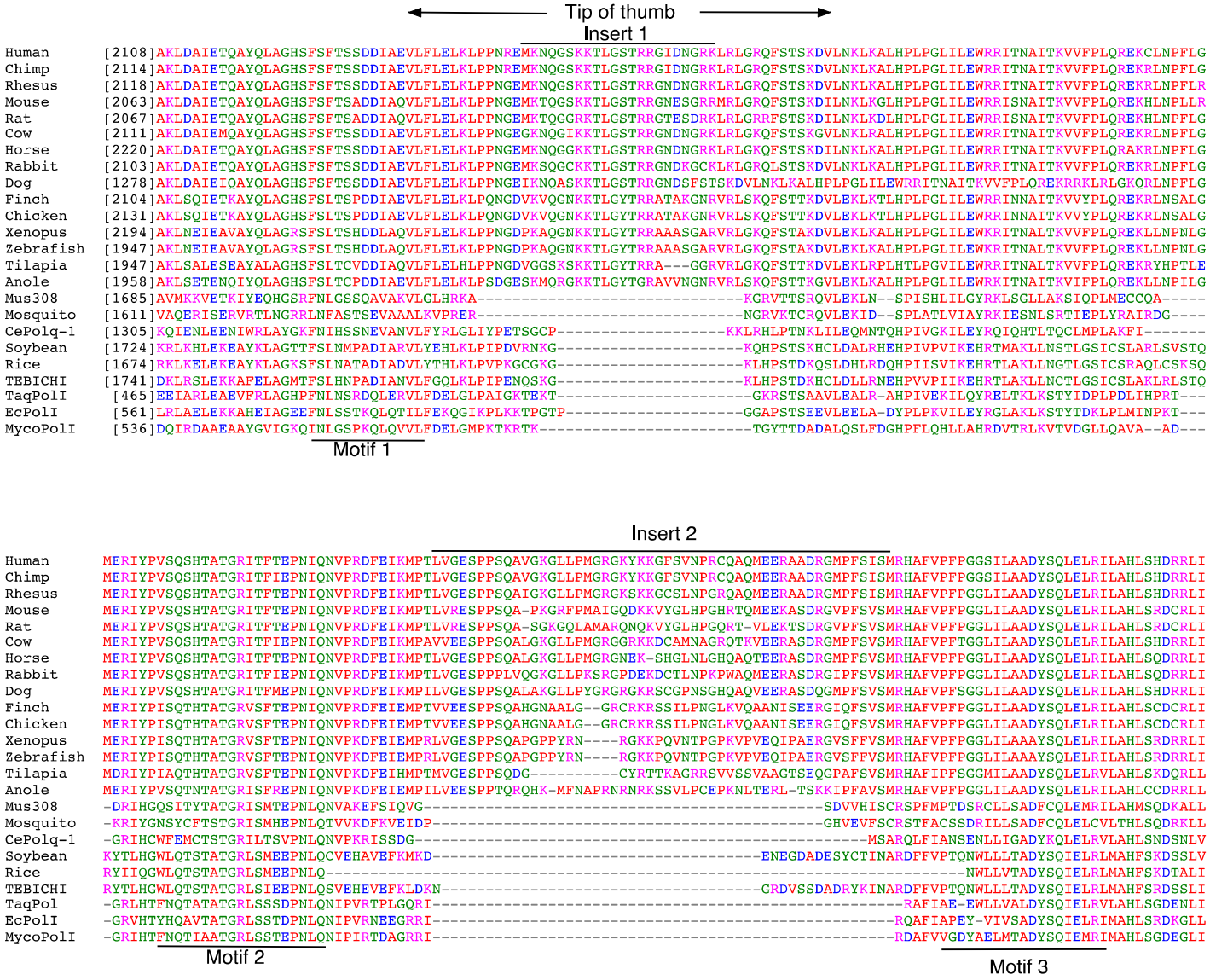
Genes with similarity to *POLQ* and *Mus308* are present in protists, plants, as well as eukaryotes (**Figure 1**), but not in yeast or other fungi. It remains to be determined whether all POLQ-like genes are true orthologs, as it appears that there are functional differences between species. Mutations in the *polq-1* gene of *C. elegans* confer sensitivity to ICL-inducing agents (Muzzini et al., 2008). TEBICHI, a POLQ-like protein found in *Arabidopsis*, is important for normal plant development and contributes to alleviating DNA replication stress and perhaps recombination. *Tebichi* mutant plants are sensitive to both the ICL-inducing agent mitomycin C (MMC) and to the monofunctional alkylating agent methyl methane sulfonate (MMS) (Inagaki et al., 2009; Inagaki et al., 2006).

The availability of numerous high quality protein sequences allows direct comparison by primary alignment, revealing several evolutionary patterns in the POLQ family. The length between the core polymerase and helicase-like domains of POLQ is about 1500–1600 amino acids in vertebrates, but only about 800 residues in plant homologs and 800–1000 amino acids in invertebrate animals (**Table 1A**). Further, the positions of the polymerase domain "inserts" are conserved within the entire POLQ family, but the inserts are much shorter in the non-vertebrate family members (**Table 1B**). Expansion of the size of the central domain and the inserts apparently arose contemporaneously in a common vertebrate ancestor. The variations in the central domain and polymerase domain insert lengths may account for some of the differences in damage sensitivities associated with defects in POLQ family genes in different organisms (**Table 1A and B, Figure 2C**). Recently, a report found three RAD51 interaction sites within the central domain of POLQ and the authors suggest that POLQ interacts with RAD51 to suppress homologous recombination (Ceccaldi et al., 2015). This serves as the first suggestion that the central domain may have an important role in the function of POLQ.

Two other genes in multicellular eukaryotes show significant and intriguing relationships with POLQ. HELQ (formerly HEL308) is a DNA repair helicase with amino acid sequence similarity to the helicase-like domain of POLQ. HELQ homologs are found in both animals (Marini and Wood, 2002) and archaea (Woodman and Bolt, 2009), but not in fungi or bacteria. POLN is a 900 amino acid protein in human cells, harboring an A-family DNA polymerase domain related to that of Mus308 and POLQ (Marini et al., 2003; Takata et al., 2010; Takata et al., 2006). In contrast with POLQ, the phylogenetic distribution of POLN appears to be limited to the deuterostome lineage of eukaryotes, including vertebrates. Phylogenetic analysis suggests that POLN is no more related to POLQ than it is to the bacterial pol I polymerases (**Figure 2C**).

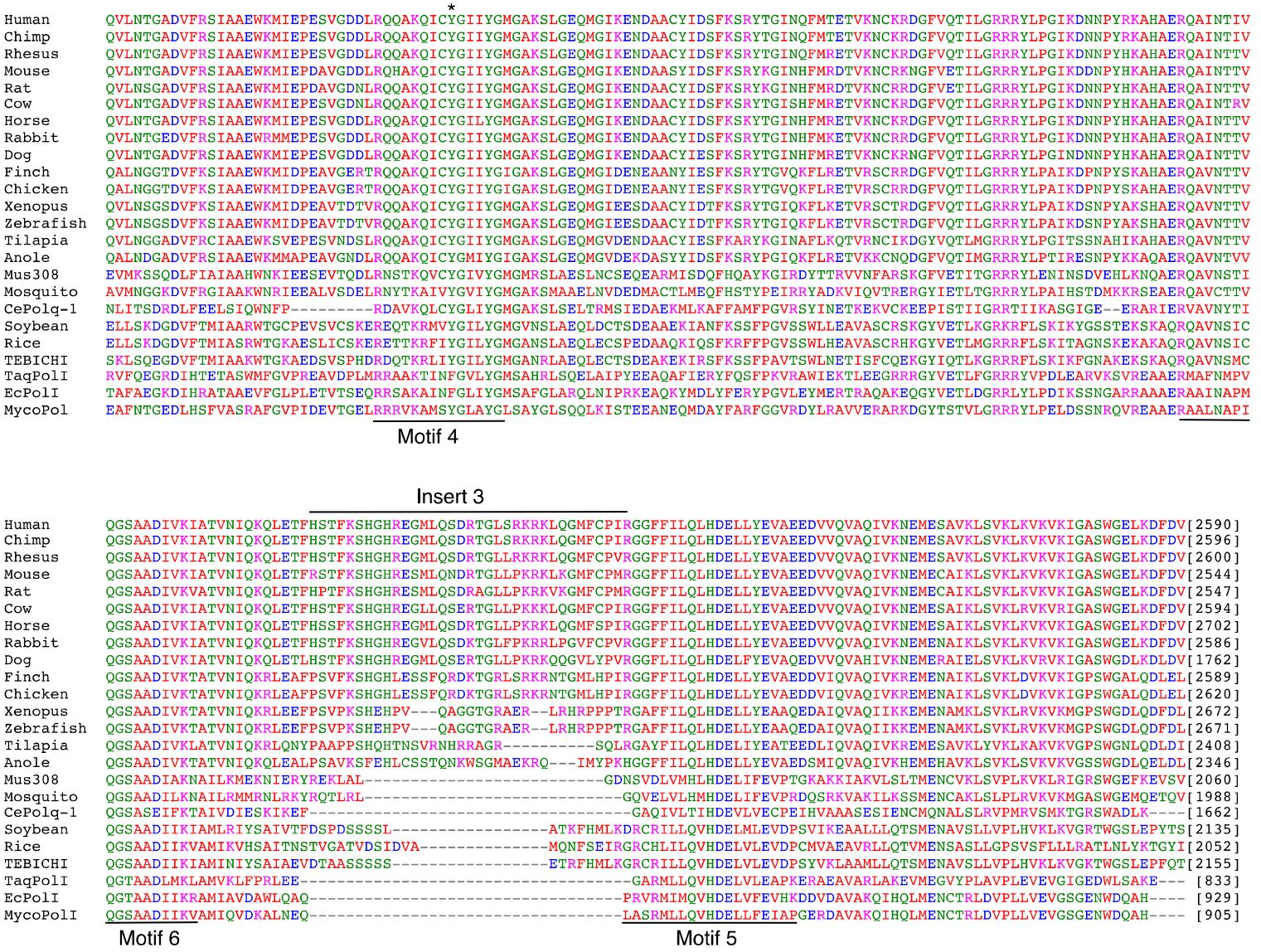
A


1819
Inactive DnaQ-like 3'->5' exonuclease domain
2090
1792
Chaos1
Exo1
Exo2
1859
1898
1918
1940
~Central Domain
2091
2125
Motif 1
2143
Insert 1
2177
2226
Motif 2
2243
2256
Insert 2
2288
Putative 5'-dRP lyase activity domain
2500
2308
Motif 3
2330
2372
Motif 4
2385
2460
Motif 6
2471
2490
Insert 3
2522
Motif 5
2540
2590
Polymerase Domain


Tip of thumb
Insert 1
Human [2108]AKLDAIETQAYQLAGHSFSFTSSDDIAEVLFLELKLPPNREMKNQGSKKTLGSTRRGIDNGRKLRLGRQFSTSKDVLNKLKALHPLPGLILEWRRITNAITKVVFPLQREKCLNPFLG
Chimp [2114]AKLDAIETQAYQLAGHSFSFTSSDDIAEVLFLELKLPPNREMKNQGSKKTLGSTRRGIDNGRKLRLGRQFSTSKDVLNKLKALHPLPGLILEWRRITNAITKVVFPLQREKRLNPFLG
Rhesus [2118]AKLDAIETQAYQLAGHSFSFTSSDDIAEVLFLELKLPPNGEMKNQGSKKTLGSTRRGNDNGRKLRLGRQFSTSKDVLNKLKALHPLPGLILEWRRITNAITKVVFPLQREKRLNPFLR
Mouse [2063]AKLDAIETQAYQLAGHSFSFTSADDIAQVLFLELKLPPNGEMKTQGSKKTLGSTRRGNESGRRMRLGRQFSTSKDILNKLKGLHPLPGLILEWRRISNAITKVVFPLQREKHLNPLLR
Rat [2067]AKLDAIETQAYQLAGHSFSFTSADDIAQVLFLELKLPPNGEMKTQGGRKTLGSTRRGTESDRKLRLGRRFSTSKDILNKLKDLHPLPGLILEWRRISNAITKVVFPLQREKHLNPFLG
Cow [2111]AKLDAIEMQAYQLAGHSFSFTSSDDIAEVLFLELKLPPNGEGKNQGIKKTLGSTRRGNDNGRKLRLGKQFSTSKGVLNKLRALHPLPGLILEWRRITNAITKVVFPLQREKRLNPFLG
Horse [2220]AKLDAIETQAYQLAGHSFSFTSSDDIAEVLFLELKLPPNGEMKNQGGKKTLGSTRRGNDNGRKLRLGKQFSTSKDILNKLKALHPLPGLILEWRRITNAITKVVFPLQRAKRLNPFLG
Rabbit [2103]AKLDAIETQAYQLAGHSFSFTSSDDIAEVLFLELKLPPNGEMKSQGCKKTLGSTRRGNDKGCKLKLGRQLSTSKDVLNKLKALHPLPGLILEWRRITNAITKVVFPLQREKRLNPFLG
Dog [1278]AKLDAIEIQAYQLAGHSFSFTSSDDIAEVLFLELKLPPNGEIKNQASKKTLGSTRRGNDSFSTSKDVLNKLKALHPLPGLILEWRRITNAITKVVFPLQREKRRKLRLGKQRLNPFLG
Finch [2104]AKLSQIETKAYQLAGHSFSLTSPDDIAEVLFLELKLPQNGDVKVQGNKKTLGYTRRATAKGNRVRLSKQFSTTKDVLEKLKTLHPLPGLILEWRRINNAITKVVYPLQREKRLNSALG
Chicken [2131]AKLSQIETKAYQLAGHSFSLTSPDDIAEVLFLELKLPQNGDVKVQGNKKTLGYTRRATAKGNRVRLSKQFSTTKDVLEKLKTLHPLPGLILEWRRINNAITKVVYPLQREKRLNSALG
Xenopus [2194]AKLNEIEAVAYQLAGRSFSLTSHDDLAQVLFLELKLPPNGDPKAQGNKKTLGYTRRAAASGARVRLGKQFSTAKDVLEKLKALHPLPGLILEWKRITNALTKVVFPLQREKLLNPNLG
Zebrafish [1947]AKLNEIEAVAYQLAGRSFSLTSHDDLAQVLFLELKLPPNGDPKAQGNKKTLGYTRRAAASGARVRLGKQFSTAKDVLEKLKALHPLPGLILEWKRITNALTKVVFPLQREKLLNPNLG
Tilapia [1947]AKLSALESEAYALAGHSFSLTCVDDIAQVLFLELHLPPNGDVGGSKSKKTLGYTRRA---GGRVRLGKQFSTTKDVLEKLRPLHTLPGVILEWRRITNALTKVVFPLQREKRYHPTLE
Anole [1958]AKLSETENQIYQLAGHSFSLTCPDDIAEVLFLELKLPSDGESKMQRGKKTLGYTGRAVVNGNRVRLSKQFSTTKGVLEKLKALHPLPGLILEWRKITNAITKVIFPLQREKLLNPILG
Mus308 [1685]AVMKKVETKIYEQHGSRFNLGSSQAVAKVLGLHRKA----------KGRVTTSRQVLEKLN--SPISHLILGYRKLSGLLAKSIQPLMECCQA-----
Mosquito [1611]VAQERISERVRTLNGRRLNFASTSEVAAALKVPRER----------NGRVKTCRQVLEKID--SPLATLVIAYRKIESNLSRTIEPLYRAIRDG----
CePolq-1 [1305]KQIENLEENIWRLAYGKFNIHSSNEVANVLFYRLGLIYPETSGCP----------KKLRHLPTNKLILEQMNTQHPIVGKILEYRQIQHTLTQCLMPLAKFI-------
Soybean [1724]KRLKHLEKEAYKLAGTTFSLNMPADIARVLYEHLKLPIPDVRNKG----------KQHPSTSKHCLDALRHEHPIVPVIKEHRTMAKLLNSTLGSICSLARLSVSTQ
Rice [1674]RKLKELEKEAYKLAGKSFSLNATADIADVLYTHLKLPVPKGCGKG----------KLHPSTDKQSLDHLRDQHPIISVIKEHRTLAKLLNGTLGSICSRAQLCSKSQ
TEBICHI [1741]DKLRSLEKKAFELAGMTFSLHNPADIANVLFGQLKLPIPENQSKG----------KLHPSTDKHCLDLLRNEHPVVPIIKEHRTLAKLLNCTLGSICSLAKLRLSTQ
TaqPolI [465]EEIARLEAEVFRLAGHPFNLNSRDQLERVLFDELGLPAIGKTEKT----------GKRSTSAAVLEALR-AHPIVEKILQYRELTKLKSTYIDPLPDLIHPRT---
EcPolI [561]LRLAELEKKAHEIAGEEFNLSSTKQLQTILFEKQGIKPLKKTPGTP----------GGAPSTSEEVLEELA-DYPLPKVILEYRGLAKLKSTYTDKLPLMINPKT---
MycoPolI [536]DQIRDAAEAAYGVIGKQINLGSPKQLQVVLFDELGMPKTKRTK----------TGYTTDADALQSLFDGHPFLQHLLAHRDVTRLKVTVDGLLQAVA--AD---
Motif 1

Insert 2
Human MERIYPVSQSHTATGRITFTEPNIQNVPRDFEIKMPTLVGESPPSQAVGKGLLPMGRGKYKKGFSVNPRCQAQMEERAADRGMPFSISMRHAFVPFPGGSILAADYSQLELRILAHLSHDRRLI
Chimp MERIYPVSQSHTATGRITFIEPNIQNVPRDFEIKMPTLVGESPPSQAVGKGLLPMGRGKYKKGFSVNPRCQAQMEERAADRGMPFSISMRHAFVPFPGGSILAADYSQLELRILAHLSHDRRLI
Rhesus MERIYPVSQSHTATGRITFTEPNIQNVPRDFEIKMPTLVGESPPSQAIGKGLLPMGRGKSKKGCSLNPGRQAQMEERAADRGMPFSISMRHAFVPFPGGLILAADYSQLELRILAHLSHDRRLI
Mouse MERIYPVSQSHTATGRITFTEPNIQNVPRDFEIKMPTLVRESPPSQA-PKGRFPMAIGQDKKVYGLHPGHRTQMEEKASDRGVPFSVSMRHAFVPFPGGLILAADYSQLELRILAHLSRDCRLI
Rat MERIYPVSQSHTATGRITFTEPNIQNVPRDFEIKMPTLVRESPPSQA-SGKGQLAMARQNQKVYGLHPGQRT-VLEKTSDRGVPFSVSMRHAFVPFPGGLILAADYSQLELRILAHLSRDCRLI
Cow MERIYPVSQSHTATGRITFIEPNIQNVPRDFEIKMPAVVEESPPSQALGKGLLPMGRGGRKKDCAMNAGRQTKVEERASDRGMPFSVSMRHAFVPFTGGLILAADYSQLELRILAHLSHDRRLI
Horse MERIYPVSQSHTATGRITFTEPNIQNVPRDFEIKMPTLVGESPPSQALGKGLLPMGRGNEK-SHGLNLGHQAQTEERASDRGMPFSVSMRHAFVPFPGGLILAADYSQLELRILAHLSQDRRLI
Rabbit MERIYPVSQSHTATGRITFIEPNIQNVPRDFEIKMPTLVGESPPPLVQGKGLLPKSRGPDEKDCTLNPKPWAQMEERASDRGIPFSVSMRHAFVPFPGGLILAADYSQLELRILAHLSQDRRLI
Dog MERIYPVSQSHTATGRITFMEPNIQNVPRDFEIKMPILVGESPPSQALAKGLLPYGRGRGKRSCGPNSGHQAQVEERASDQGMPFSVSMRHAFVPFSGGLILAADYSQLELRILAHLSHDRRLI
Finch MERIYPISQTHTATGRVSFTEPNIQNVPRDFEIEMPTVVEESPPSQAHGNAALG--GRCKRRSSILPNGLKVQAANISEERGIQFSVSMRHAFVPFPGGLILAADYSQLELRILAHLSCDCRLI
Chicken MERIYPISQTHTATGRVSFTEPNIQNVPRDFEIEMPTVVEESPPSQAHGNAALG--GRCKRRSSILPNGLKVQAANISEERGIQFSVSMRHAFVPFPGGLILAADYSQLELRILAHLSCDCRLI
Xenopus MERIYPISQTHTATGRVSFTEPNIQNVPKDFEIEMPRLVGESPPSQAPGPPYRN----RGKKPQVNTPGPKVPVEQIPAERGVSFFVSMRHAFVPFPGGLILAAAYSQLELRILAHLSRDRRLI
Zebrafish MERIYPISQTHTATGRVSFTEPNIQNVPKDFEIEMPRLVGESPPSQAPGPPYRN----RGKKPQVNTPGPKVPVEQIPAERGVSFFVSMRHAFVPFPGGLILAAAYSQLELRILAHLSRDRRLI
Tilapia MDRIYPIAQTHTATGRVSFTEPNIQNVPKDFEIHMPTMVGESPPSQDG----------CYRTTKAGRRSVVSSVAAGTSEQGPAFSVSMRHAFIPFSGGMILAADYSQLELRVLAHLSKDQRLL
Anole MERIYPVSQTNTATGRISFREPNIQNVPKDFEIEMPILVEESPPTQRQHK-MFNAPRNRNRKSSVLPCEPKNLTERL-TSKKIPFAVSMRHAFVPFPGGLILAADYSQLELRILAHLCCDRRLL
Mus308 -DRIHGQSITYTATGRISMTEPNLQNVAKEFSIQVG----------SDVVHISCRSPFMPTDSRCLLSADFCQLEMRILAHMSQDKALL
Mosquito -KRIYGNSYCFTSTGRISMHEPNLQTVVKDFKVEIDP----------GHVEVFSCRSTFACSSDRILLSADFCQLELCVLTHLSQDRKLL
CePolq-1 -GRIHCWFEMCTSTGRILTSVPNLQNVPKRISSDG----------MSARQLFIANSENLLIGADYKQLELRVLAHLSNDSNLV
Soybean KYTLHGWLQTSTATGRLSMEEPNLQCVEHAVEFKMKD----------ENEGDADESYCTINARDFFVPTQNWLLLTADYSQIELRLMAHFSKDSSLV
Rice RYIIQGWLQTSTATGRLSMEEPNLQ----------NWLLVTADYSQIELRLMAHFSKDTALI
TEBICHI RYTLHGWLQTSTATGRLSIEEPNLQSVEHEVEFKLDKN----------GRDVSSDADRYKINARDFFVPTQNWLLLTADYSQIELRLMAHFSRDSSLI
TaqPol -GRLHTFNQTATATGRLSSSDPNLQNIPVRTPLGQRI----------RAFIAE-EWLLVALDYSQIELRVLAHLSGDENLI
EcPolI -GRVHTYHQAVTATGRLSSTDPNLQNIPVRNEEGRRI----------RQAFIAPEY-VIVSADYSQIELRIMAHLSRDKGLL
MycoPolI -GRIHTFNQTIAATGRLSSTEPNLQNIPIRTDAGRRI----------RDAFVVGDYAELMTADYSQIEMRIMAHLSGDEGLI
Motif 2
Motif 3

B


Human QVLNTGADVFRSIAAEWKMIEPESVGDDLRQQAKQICYGIIYGMGAKSLGEQMGIKENDAACYIDSFKSRYTGINQFMTETVKNCKRDGFVQTILGRRRYLPGIKDNNPYRKAHAERQAINTIV
Chimp QVLNTGADVFRSIAAEWKMIEPESVGDDLRQQAKQICYGIIYGMGAKSLGEQMGIKENDAACYIDSFKSRYTGINQFMTETVKNCKRDGFVQTILGRRRYLPGIKDNNPYRKAHAERQAINTIV
Rhesus QVLNTGADVFRSIAAEWKMIEPESVGDDLRQQAKQICYGIIYGMGAKSLGEQMGIKENDAACYIDSFKSRYTGINQFMRETVKNCKRDGFVQTILGRRRYLPGIKDNNPYHKAHAERQAINTTV
Mouse QVLNTGADVFRSIAAEWKMIEPDAVGDDLRQHAKQICYGIIYGMGAKSLGEQMGIKENDAASYIDSFKSRYKGINHFMRDTVKNCRKNGFVETILGRRRYLPGIKDDNPYHKAHAERQAINTTV
Rat QVLNSGADVFRSIAAEWKMIEPDAVGDNLRQQAKQICYGIIYGMGAKSLGEQMGIKENDAACYIDSFKSRYKGINHFMRDTVKNCRRDGFVETILGRRRYLPGIKDNNPYHKAHAERQAINTTV
Cow QVLNTGADVFRSIAAEWKMIEPESVGDNLRQQAKQICYGIIYGMGAKSLGEQMGIKENDAACYIDSFKSRYTGISHFMRETVKNCKRDGFVQTILGRRRYLPGIKDNNPYHKAHAERQAINTRV
Horse QVLNTGADVFRSIAAEWKMIEPESVGDDLRQQAKQICYGILYGMGAKSLGEQMGIKENDAACYIDSFKSKYTGINHFMRETVKNCKRDGFVQTILGRRRYLPGIKDNNPYHKAHAERQAINTTV
Rabbit QVLNTGEDVFRSIAAEWRMMEPESVGDDLRQQAKQICYGIIYGMGAKSLGEQMGIKENDAACYIDSFKSRYTGINHFMKETVKNCRRDGFVQTILGRRRYLPGIKDNNPYHKAHAERQAINTTV
Dog QVLNTGADVFRSIAAEWKMIEPESVGDDLRQQAKQICYGIIYGMGAKSLGEQMGIKENDAACYIESFKSRYTGINHFMRETVKNCKRNGFVQTILGRRRYLPGINDNNPYHKAHAERQAINTTV
Finch QALNGGTDVFKSIAAEWKMIDPEAVGERTRQQAKQICYGIIYGIGAKSLGEQMGIDENEAANYIESFKSRYTGVQKFLRETVRSCRRDGFVQTILGRRRYLPAIKDPNPYSKAHAERQAVNTTV
Chicken QALNGGTDVFKSIAAEWKMIDPEAVGERTRQQAKQICYGIIYGIGAKSLGEQMGIDENEAANYIESFKSRYTGVQKFLRETVRSCRRDGFVQTILGRRRYLPAIKDPNPYSKAHAERQAVNTTV
Xenopus QVLNSGSDVFKSIAAEWKMIDPEAVTDTVRQQAKQICYGIIYGMGAKSLGEQMGIEESDAACYIDTFKSRYTGIQKFLKETVRSCTRDGFVQTILGRRRYLPAIKDSNPYAKSHAERQAVNTTV
Zebrafish QVLNSGSDVFKSIAAEWKMIDPEAVTDTVRQQAKQICYGIIYGMGAKSLGEQMGIEESDAACYIDTFKSRYTGIQKFLKETVRSCTRDGFVQTILGRRRYLPAIKDSNPYAKSHAERQAVNTTV
Tilapia QVLNGGADVFRCIAAEWKSVEPESVNDSLRQQAKQICYGIIYGMGAKSLGEQMGVDENDAACYIESFKARYKGINAFLKQTVRNCIKDGYVQTLMGRRRYLPGITSSNAHIKAHAERQAVNTTV
Anole QALNDGADVFRSIAAEWKMMAPEAVGNDLRQQAKQICYGMIYGIGAKSLGEQMGVDEKDASYYIDSFKSRYPGIQKFLRETVKKCNQDGFVQTIMGRRRYLPTIRESNPYKKAQAERQAVNTVV
Mus308 EVMKSSQDLFIAIAAHWNKIEESEVTQDLRNSTKQVCYGIVYGMGMRSLAESLNCSEQEARMISDQFHQAYKGIRDYTTRVVNFARSKGFVETITGRRRYLENINSDVEHLKNQAERQAVNSTI
Mosquito AVMNGGKDVFRGIAAKWNRIEEALVSDELRNYTKAIVYGVIYGMGAKSMAAELNVDEDMACTLMEQFHSTYPEIRRYADKVIQVTRERGYIETLTGRRRYLPAIHSTDMKKRSEAERQAVCTTV
CePolq-1 NLITSDRDLFEELSIQWNFP---------RDAVKQLCYGLIYGMGAKSLSELTRMSIEDAEKMLKAFFAMFPGVRSYINETKEKVCKEEPISTIIGRRTIIKASGIGE--ERARIERVAVNYTI
Soybean ELLSKDGDVFTMIAARWTGCPEVSVCSKEREQTKRMVYGILYGMGVNSLAEQLDCTSDEAAEKIANFKSSFPGVSSWLLEAVASCRSKGYVETLKGRKRFLSKIKYGSSTEKSKAQRQAVNSIC
Rice ELLSKDGDVFTMIASRWTGKAESLICSKERETTKRFIYGILYGMGANSLAEQLECSPEDAAQKIQSFKRFFPGVSSWLHEAVASCRHKGYVETLMGRRRFLSKITAGNSKEKAKAQRQAVNSIC
TEBICHI SKLSQEGDVFTMIAAKWTGKAEDSVSPHDRDQTKRLIYGILYGMGANRLAEQLECTSDEAKEKIRSFKSSFPAVTSWLNETISFCQEKGYIQTLKGRRRFLSKIKFGNAKEKSKAQRQAVNSMC
TaqPolI RVFQEGRDIHTETASWMFGVPREAVDPLMRRAAKTINFGVLYGMSAHRLSQELAIPYEEAQAFIERYFQSFPKVRAWIEKTLEEGRRRGYVETLFGRRRYVPDLEARVKSVREAAERMAFNMPV
EcPolI TAFAEGKDIHRATAAEVFGLPLETVTSEQRRSAKAINFGLIYGMSAFGLARQLNIPRKEAQKYMDLYFERYPGVLEYMERTRAQAKEQGYVETLDGRRLYLPDIKSSNGARRAAAERAAINAPM
MycoPol EAFNTGEDLHSFVASRAFGVPIDEVTGELRRRVKAMSYGLAYGLSAYGLSQQLKISTEEANEQMDAYFARFGGVRDYLRAVVERARKDGYTSTVLGRRRYLPELDSSNRQVREAAERAALNAPI
Motif 4
Insert 3
Human QGSAADIVKIATVNIQKQLETFHSTFKSHGHREGMLQSDRTGLSRKRKLQGMFCPIRGGFFILQLHDELLYEVAEEDVVQVAQIVKNEMESAVKLSVKLKVKVKIGASWGELKDFDV[2590]
Chimp QGSAADIVKIATVNIQKQLETFHSTFKSHGHREGMLQSDRTGLSRKRKLQGMFCPIRGGFFILQLHDELLYEVAEEDVVQVAQIVKNEMESAVKLSVKLKVKVKIGASWGELKDFDV[2596]
Rhesus QGSAADIVKIATVNIQKQLETFHSTFKSHGHREGMLQSDRTGLLRKRKLQGMFCPVRGGFFILQLHDELLYEVAEEDVVQVAQIVKNEMESAVKLSVKLKVKVKIGASWGELKDFDV[2600]
Mouse QGSAADIVKIATVNIQKQLETFRSTFKSHGHRESMLQNDRTGLLPKRKLKGMFCPMRGGFFILQLHDELLYEVAEEDVVQVAQIVKNEMECAIKLSVKLKVKVKIGASWGELKDFDV[2544]
Rat QGSAADIVKVATVNIQKQLETFHPTFKSHGHRESMLQSDRAGLLPKRKVKGMFCPMRGGFFILQLHDELLYEVAEEDVVQVAQIVKNEMECAIKLSVKLKVKVKIGASWGELKDFDV[2547]
Cow QGSAADIVKIATVNIQKQLETFHSTFKSHGHREGLLQSERTGLLPKKKLQGMFCPIRGGFFILQLHDELLYEVAEEDVVQVAQIVKNEMESAIKLSVKLRVKVRIGASWGELKDFDV[2594]
Horse QGSAADIVKIATVNIQKQLETFHSSFKSHGHREGMLQSDRTGLLPKRKLQGMFSPIRGGFFILQLHDELLYEVAEEDVVQVAQIVKNEMESAIKLSVKLKVKVKIGASWGELKDFDV[2702]
Rabbit QGSAADIVKIATVNIQKQLETFHSTFKSHGHREGVLQSDKTGLFPKRRLPGVFCPVRGGFFILQLHDELLYEVAEEDVVQVAQIVKNEMENAIKLSVKLKVKVKIGASWGELKDFDV[2586]
Dog QGSAADIVKIATVNIQKQLETLHSTFKSHGHREGMLQSERTGLLPKRKQQGVLYPVRGGFLILQLHDELFYEVAQEDVVQVAHIVKNEMERAIELSVKLKVRVKIGASWGDLKDLDV[1762]
Finch QGSAADIVKTATVNIQKRLEAFPSVFKSHGHLESSFQRDKTGRLSRKRNTGMLHPIRGGFFILQLHDELLYEVAEDDVIQVAQIVKREMENAIKLSVKLDVKVKIGPSWGALQDLEL[2589]
Chicken QGSAADIVKTATVNIQKRLEAFPSVFKSHGHLESSFQRDKTGRLSRKRNTGMLHPIRGGFFILQLHDELLYEVAEDDVIQVAQIVKREMENAIKLSVKLDVKVKIGPSWGALQDLEL[2620]
Xenopus QGSAADIVKTATVNIQKRLEEFPSVPKSHEHPV---QAGGTGRAER--LRHRPPPTRGAFFILQLHDELLYEAAQEDAIQVAQIIKKEMENAMKLSVKLRVKVKMGPSWGDLQDFDL[2672]
Zebrafish QGSAADIVKTATVNIQKRLEEFPSVPKSHEHPV---QAGGTGRAER--LRHRPPPTRGAFFILQLHDELLYEAAQEDAIQVAQIIKKEMENAMKLSVKLRVKVKMGPSWGDLQDFDL[2671]
Tilapia QGSAADIVKLATVNIQKRLQNYPAAPPSHQHTNSVRNHRRAGR----------SQLRGAYFILQLHDELIYEATEEDLIQVAQIVKREMESAVKLYVKLKAKVKVGPSWGNLQDLDI[2408]
Anole QGSAADIVKTATVNIQKQLEALPSAVKSFEHLCSSTQNKWSGMAEKRQ---IMYPKHGGFFILQLHDELLYEVAEDSMIQVAQIVKHEMEHAVKLSVKLSVKVKVGSSWGDLQELDL[2346]
Mus308 QGSAADIAKNAILKMEKNIERYREKLAL-------------------------GDNSVDLVMHLHDELIFEVPTGKAKKIAKVLSLTMENCVKLSVPLKVKLRIGRSWGEFKEVSV[2060]
Mosquito QGSAADILKNAILRMMRNLRKYRQTLRL----------------------------GQVELVLHMHDELIFEVPRDQSRKVAKILKSSMENCAKLSLPLRVKVKMGASWGEMQETQV[1988]
CePolq-1 QGSASEIFKTAIVDIESKIKEF-------------------------------GAQIVLTIHDEVLVECPEIHVAAASESIENCMQNALSLRVPMRVSMKTGRSWADLK----[1662]
Soybean QGSAADIIKIAMLRIYSAIVTFDSPDSSSSL-----------------ATKFHMLKDRCRILLQVHDELMLEVDPSVIKEAALLLQTSMENAVSLLVPLHVKLKVGRTWGSLEPYTS[2135]
Rice QGSAADIIKVAMIKVHSAITNSTVGATVDSIDVA--------------MQNFSEIRGRCHLIIQVHDELVLEVDPCMVAEAVRLLQTVMENSASLLGPSVSFLLLRATLNLYKTGYI[2052]
TEBICHI QGSAADIIKIAMINIYSAIAEVDTAASSSSS-----------------ETRFHMLKGRCRILLQVHDELVLEVDPSYVKLAAMLLQTSMENAVSLLVPLHVKLKVGKTWGSLEPFQT[2155]
TaqPolI QGTAADLMKLAMVKLFPRLEE-----------------------------------GARMLLQVHDELVLEAPKERAEAVARLAKEVMEGVYPLAVPLEVEVGIGEDWLSAKE--- [833]
EcPolI QGTAADIIKRAMIAVDAWLQAQ-----------------------------------PRVRMIMQVHDELVFEVHKDDVDAVAKQIHQLMENCTRLDVPLLVEVGSGENWDQAH---- [929]
MycoPolI QGSAADIIKVAMIQVDKALNEQ-----------------------------------LASRMLLQVHDELLFEIAPGERDAVAKQIHQLMENCTRLDVPLLVEVGSGENWDQAH---- [905]
Motif 6
Motif 5

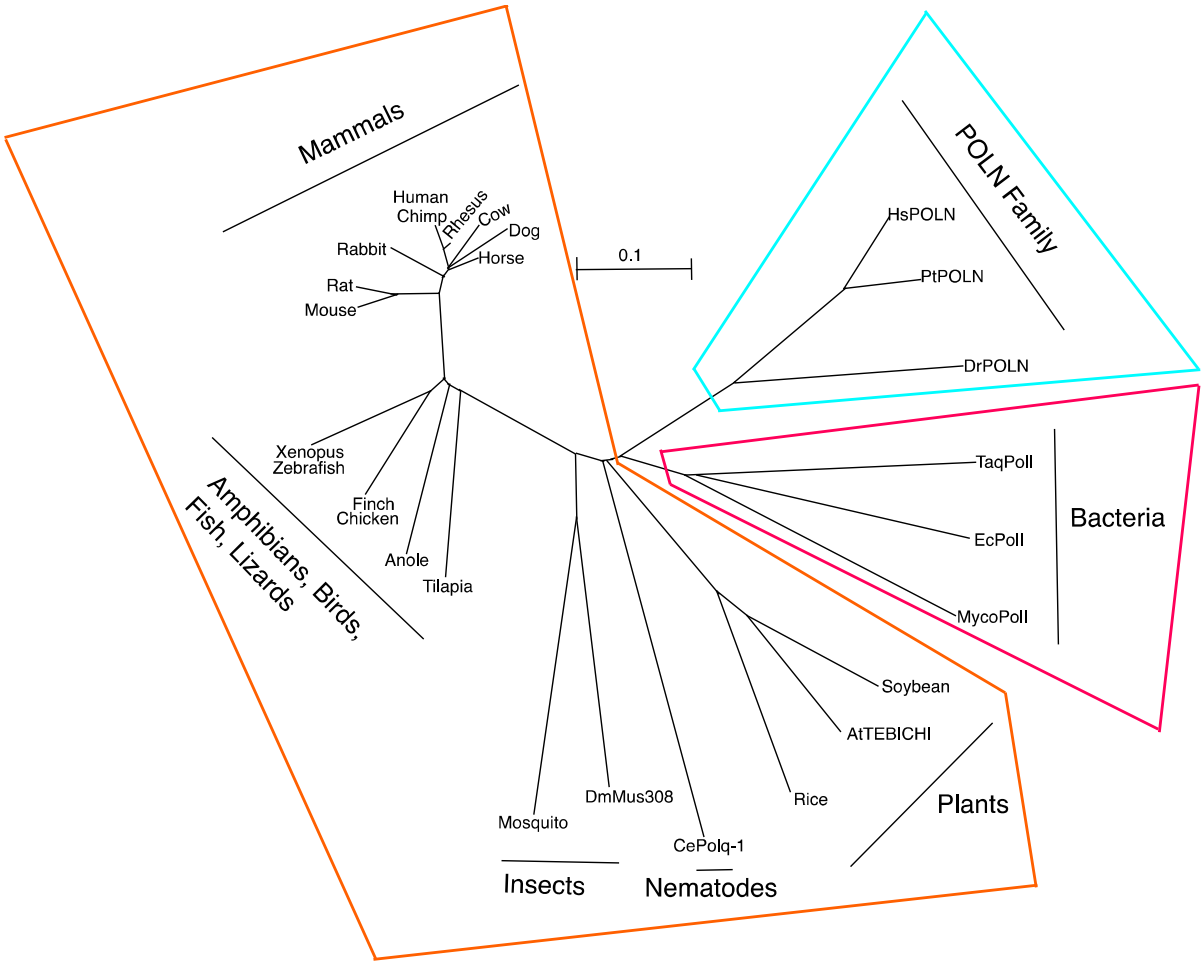


Figure 2. Map and alignment of POLQ polymerase domain. (A) The C-terminal portion of POLQ (amino acids 1792-2590) necessary for polymerase activity is displayed. The core polymerase domain of POLQ is displayed in blue. Conserved A-family DNA polymerase motifs are noted in black. The unique sequence inserts found in POLQ are shown in yellow. An inactive DnaQ-like 3'→5' exonuclease domain (including insertion loops shaded in orange), Chaos1 mutation (Ser1932 in mouse/Ser1977 in human), and the 5'-dRP lyase activity of POLQ mapped to the polymerase domain are noted. An asterisk notes the tyrosine residue 2337 in motif 4 of the polymerase domain that facilitates ddNTP incorporation. The location of the serine to proline mutation found in the *Chaos1* mutant mouse is noted in the central domain. (B) Sequence alignment of POLQ family members

constructed in MacVector. Amino acid numbers for each species in the alignment are notated in brackets. (C). The dendogram illustrating evolutionary relatedness of POLQ gene family members was created using SeaView. POLQ, POLN, and bacterial pol I homologs are shown in orange, cyan, and magenta outlined groups, respectively.

MycoPolI: *Mycobacterium tuberculosis* pol I; EcPolI: *Escherichia coli* pol I; TaqPolI: *Thermus aquaticus* pol I; TEBICHI: *Arabidopsis thaliana* TEBICHI; Rice: *Oryza sativa* pol I; Soybean: *Glycine max* POLQ-like; CePolq-1: *Caenorhabditis elegans* Polq-1; Mosquito: *Aedes aegypti* POLQ; Mus308: *Drosophlia melanogaster* Mus308; Anole: *Anolis carolinensis* POLQ-like; Tilapia: *Oreochromis niloticus* POLQ; Zebrafish: *Danio rerio* POLQ; Xenopus: *Xenopus tropicalis* POLQ; Chicken: *Gallus gallus* POLQ; Finch: *Taeniopygia guttata* POLQ; Dog: *Canis lupus* POLQ; Rabbit: *Oryctolagus cuniculus*; Horse: *Equus caballus* POLQ; Cow: *Bos taurus* POLQ; Rat: *Rattus norvegicus* POLQ; Mouse: *Mus musculus* POLQ; Rhesus: *Macaca mulatta* POLQ; Chimp: *Pan troglodytes* POLQ; Human: *Homo sapiens* POLQ; HsPOLN: *Homo sapiens* POLN; PtPOLN: *Pan troglodytes* POLN; MmPOLN: *Mus musculus* POLN; DrPOLN: *Danio rerio* POLN.

The *scale bar* indicates a distance of 0.1 amino acid substitution per site.

A

Distance between core catalytic domains		
Species	Length (aa)	
Human	1565	
Chimp	1596	
Mouse	1552	
Cow	1593	
Horse	1599	Vertebrates
Chicken	1610	
Finch	1581	
Xenopus	1557	
Zebrafish	1457	
Tilapia	1271	
Anole	1364	
Pea Aphid	905	
Western Mite	1018	
Fruit Fly	1042	Invertebrates
Mosquito	801	
C. elegans	888	
Arabidopsis	806	
Soybean	817	Plants

B

Length of POLQ polymerase domain inserts (aa)				
Species	Insert 1	Insert 2	Insert 3	
Human	22	52	33	
Chimp	22	52	33	
Mouse	22	52	33	
Chicken	22	52	33	Vertebrates
Dog	22	52	33	
Xenopus	22	48	28	
Zebrafish	19	51	24	
Tilapia	19	42	25	
Fruit Fly	4	9	9	
Mosquito	0	10	7	Invertebrates
C. elegans	3	9	0	
Soybean	5	18	13	
Rice	5	0	16	Plants
Arabidopsis	5	17	13	

Table 1. Conservation of POLQ functional interdomain distance and polymerase domain inserts in evolution. Length of (**A**) the amino acid separation between the core catalytic domains of the helicase-like and polymerase domains and (**B**) polymerase domain inserts in POLQ gene family members. All lengths are in amino acids.

Low fidelity and translesion synthesis activities of POLQ

Human POLQ has been purified as a recombinant protein from a baculovirus-infected insect cell expression system (Seki et al., 2003; Seki et al., 2004). The protein is active on substrates including oligonucleotide primer-templates, hairpin primer-templates, activated calf thymus DNA and poly (dA)-oligo (dT) (Seki et al., 2003; Seki et al., 2004). The activity of recombinant POLQ is relatively resistant to aphidicolin, a potent inhibitor of eukaryotic replicative DNA polymerases (pols α,δ,ε) (Cheng and Kuchta, 1993; Dresler et al., 1988; Seki et al., 2003; Sheaff et al., 1991). POLQ is sensitive to dideoxynucleoside

triphosphate (ddNTP), consistent with the presence of a tyrosine residue in motif 4 of family-A DNA polymerases that facilitates ddNTP incorporation (**Figure 2B**) (Doublié et al., 1998; Seki et al., 2003). Sequence alignment shows that POLQ contains an inactive DnaQ-like 3'→5' exonuclease that lacks conservation in the key HxAxxD ExoIIIε metal binding motif. Purified POLQ lacks a 3'→5' exonuclease activity, which is a common feature of low fidelity DNA polymerases (**Figure 2A**). Primer extension assays to measure DNA polymerase fidelity, found that purified POLQ frequently misincorporated a G or T across from a T in the template (Seki et al., 2003; Seki et al., 2004). In a gap-filling assay to measure mutation frequency, human POLQ generated single base errors at a 10 to 100-fold higher rate than the other A family DNA polymerase members (γ, ν) (Arana et al., 2008). POLQ adds single nucleotides to homopolymeric runs at particularly high rates during gap filling, exceeding 1% in certain sequence contexts. The enzyme generates single base substitutions at an average rate of 2.4×10^{-3}, comparable in rate to the inaccurate family Y human pol κ (5.8×10^{-3}). This is unusual, as the most studied low fidelity polymerases belong to the Y family (η, ι, κ, Rev1) rather than the A family of polymerases (Goodman, 2002).

Maga *et al.* purified an activity, thought to be from POLQ, from HeLa cell nuclear extracts that could efficiently utilize abasic site-containing DNA as a template. The purified preparation contained several polypeptides of about 100 kDa, one of which was recognized by an antibody raised against a fragment of recombinant POLQ (Maga et al., 2002). The active fraction exhibited relatively high fidelity of DNA polymerization, and a 3' → 5' exonuclease activity, but no ATPase or helicase activities (Maga et al., 2002). It seems likely that this preparation may have contained an active fragment of POLQ, co-purifying with a 3' → 5' nuclease activity. The identity of this exonuclease is unknown but it may be physiologically relevant to POLQ activity.

POLQ has the ability to catalyze DNA synthesis across specific forms of endogenous and exogenous DNA damage in a process called translesion DNA synthesis (TLS). A notable POLQ enzymatic trait is its ability to incorporate an A residue opposite abasic (AP) sites and extend from the incorporated nucleotide. Although many DNA polymerases tend to insert an A opposite an AP site (known as the "A" rule (Avkin et al., 2002; Strauss et al., 1982; Strauss, 2002), the highly efficient extension from the A opposite an AP site is unique to POLQ (Seki et al., 2004). For example, POLH (η) can insert a base opposite an AP site but cannot efficiently extend from the lesion, a property common to human Y family polymerases (Johnson et al., 2000; Lawrence, 2002; Masutani et al., 2000; Seki et al., 2004; Vaisman et al., 2002; Zhang et al., 2001). About 18,000 AP sites arise in each diploid mammalian cell per day by hydrolytic depurination, but insertion of an A opposite such a site would be mutagenic (Friedberg et al., 2006). Abasic sites are also generated frequently as intermediates in the base excision repair of uracil following the accidental incorporation of dUTP during genomic DNA replication (Guillet and Boiteux, 2003; Lindahl, 1993; Nakamura et al., 1998). If a DNA replication fork encounters such an AP site before it is repaired, introduction of an A would reconstitute the original undamaged template and be physiologically useful. It remains to be determined whether this AP site bypass activity is a function of POLQ *in vivo*.

POLQ cannot insert a base opposite the common UV-radiation induced cyclobutane pyrimidine dimers or 6-4 photoproducts (6-4PP). However, if pol ι (Vaisman et al., 2004) is used to incorporate bases opposite a 6-4PP, POLQ can extend the poorly matched primer-terminus (Seki et al., 2004; Seki and Wood, 2008). POLQ was also able to efficiently extend mismatched A:G, A:T, and A:C termini (Seki and Wood, 2008). The ability of POLQ to catalyze extension from minimally primed or mispaired DNA ends is quite important for its function in cells and is discussed in Chapter 2 of this thesis (Yousefzadeh et al., 2014).

Thymine glycol (Tg) is a common product of reactive oxygen species-mediated damage to DNA. POLQ and POLN can efficiently bypass both enantiomers (5*R*,5*S*) of a Tg adduct, whereas pol η extends efficiently only from the 5*R*-Tg (Kusumoto et al., 2002; Seki et al., 2004; Takata et al., 2006). Translesion synthesis bypass of a Tg residing on a shuttle vector was reduced in human MRC5 cell lines with diminished levels of POLQ (Yoon et al., 2014). This finding was clouded, in my opinion, by knowledge that siRNA suppression of POLQ can be variable and knockdown was measured by semiquantitative RT-PCR, rather than other detection methods that are more sensitive and quantitative. Furthermore, immunoblotting for POLQ was carried out using an unvalidated rabbit polyclonal antibody. The Abcam ab80906 rabbit polyclonal antibody against POLQ used by the authors is reported by the manufacturer to recognize a band between 144 and 168 kDa when blotted against placenta lysate. Notably, POLQ has a predicted molecular weight of 290 kDa. Further, the authors did not show full immunoblots thus leaving open the question of whether they actually detected the proper band for POLQ or whether it was some other non-specific band. The authors presented similar controversial immunoblotting results with a rabbit polyclonal antibody against POLN (Abnova POLN A01 H000353497-A01), thus calling additional experimental conclusions into doubt regarding the knockdown efficiency of either POLN or POLQ by siRNA. If the findings do hold true, it suggests that POLQ bypasses Tg adducts *in vivo*, through an error prone TLS process. However if POLQ cannot bypass Tg *in vivo* then *in vitro* evidence of POLQ's ability to bypass Tg could be explained by the enzyme's unique ability to utilize minimally or misprimed ends to catalyze extension (Seki and Wood, 2008; Yousefzadeh et al., 2014). Lesion bypass by POLQ could be useful in catalyzing the extension of 3' single-stranded DNA ends (See Chapter 2 of this thesis), which can be easily damaged through for example methylation, deamination, and oxidization.

The ability of POLQ to extend from mismatched, poorly matched or unmatched termini is unusual and will be understood better after high-resolution structural features of the active site are determined. The three “inserts” in the catalytic domain are unique to POLQ (**Figure 2A and B**), and some specific deletions of these inserts have been analyzed with respect to activity. Insert 1 comprises 22 amino acids between the first and second conserved motifs in the “thumb” subdomain, and its sequence is highly conserved throughout vertebrates (Seki et al., 2004). In other A-family DNA polymerases, the thumb region influences DNA binding and frameshift fidelity. Deletions in the thumb region of *E. coli* Pol I can cause errors in processivity and an increased use of misaligned primer/template complexes (Cannistraro and Taylor, 2004; Doublié et al., 1998; Minnick et al., 1996). Insert 2 is 52 amino acids long and falls between the second and third motif and has the least sequence conservation among the inserts while insert 3 is 33 amino acids that lies between the fifth and sixth motif (**Figure 2A and B**). A fragment of POLQ (residues 1792–2590) containing the pol domain and a portion of the central domain (**Figure 2A**), is active as a DNA polymerase when produced as a recombinant protein in *E. coli*. Shorter fragments are much less active, or inactive (Hogg et al., 2011; Prasad et al., 2009) indicating that this region of the central domain is required for efficient polymerase activity. The active fragment is able to bypass AP sites and Tg adducts, showing that these bypass activities are independent of the N-terminal helicase-like domain. Elimination of Insert 1 of this POLQ fragment reduced processivity of the enzyme but had little, if any, bearing on the translesion synthesis properties of the enzyme. However, removal of either inserts 2 or 3 reduced activity on undamaged DNA and eliminated the ability of POLQ to bypass AP sites or Tg adducts (Hogg et al., 2011).

The helicase-like domain of POLQ has 7 conserved motifs of the superfamily II DNA and RNA helicase family (Seki et al., 2003). The POLQ helicase-like domain exhibits some

single-stranded DNA-dependent ATPase activity, though lower in comparison to that of HELQ. No overt helicase activity of POLQ has been reported (Seki et al., 2003). It is possible that POLQ will only display helicase activity with an as yet untested substrate or alternatively, requires additional accessory factors. Another possibility is that the helicase-like domain of POLQ could function as a protein-displacement motor, in lieu of unwinding DNA. Several cellular functions have been suggested for POLQ. These are summarized in the next sections.

Hematopoiesis and immunological diversity

Some investigations have explored the possibility of a specialized immunological function for POLQ, either in somatic hypermutation (SHM) or in class switch recombination (CSR) of immunoglobulin genes. *POLQ* is expressed in hematopoietic cells, including germinal center B cells (Kawamura et al., 2004). SHM acts to introduce point mutations in the variable (V) region of antibodies to alter their affinities while CSR changes the immunoglobulin constant region. Both processes are set in motion by the AID enzyme which can deaminate cytosine to uracil in DNA (Honjo et al., 2004). Mutations may then arise from low-fidelity repair synthesis or nucleotide misincorporation opposite an AP site (Neuberger et al., 2005; Seki et al., 2005).

With respect to SHM, *Polq*$^{-/-}$ mice were previously reported to have decreased mutations of C:G sequences at SHM hotspots, whereas mutations at A:T sequences were unaffected (Masuda et al., 2005). POLH is important for A:T mutagenesis in SHM and *Polh*$^{-/-}$ *Polq*$^{-/-}$ mice do not exhibit a decrease in A:T mutations when compared to *Polh*$^{-/-}$ mice (Masuda et al., 2007; Ukai et al., 2006). POLQ extends beyond mispairs with A or T and can extend weakly from mispaired termini generated by POLH, suggesting that

POLQ and POLH might cooperate during A:T mutagenesis (Masuda et al., 2007). In an independent set of experiments, mice carrying a gene disruption that completely eliminates the POLQ enzyme had a frequency of SHM that was reduced by 60–80% (Zan et al., 2005). There was no overall change in the proportion of events at A:T and C:G positions. Finally, Martomo *et al*. found similar mutation frequencies of V genes in Peyer's patches of *Polh*$^{-/-}$ *Polq*$^{+/+}$ and *Polh*$^{-/-}$ *Polq*$^{-/-}$ mutant mice, suggesting that POLQ has only a minor role, if any, in SHM point mutagenesis (Martomo et al., 2008).

The role of POLQ in immunoglobulin (*Ig*) gene conversion, a functionally analogous process to SHM in avian species, but biochemically quite distinct, as it relies on homologous recombination rather than point mutagenesis, was investigated. The *Polq* gene was disrupted in the DT40 chicken B cell line by Yoshimura *et al*. using gene targeting (Kohzaki et al., 2010; Yoshimura et al., 2006). Strikingly, a *Poln Polq* double mutant showed a five-fold decrease in *Ig* gene conversion, the primary method of V region diversity in avian species. Further, a triple *Polh Poln Polq* mutant failed to exhibit any Ig gene conversion (Kohzaki et al., 2010).

The absence of POLQ does not seem to impair CSR in mammals as the mouse CH12F3 B-cell line (defective for *Polq*) can efficiently switch from IgM to IgA after appropriate stimulation (Li et al., 2011). *In vivo* CSR frequency was not impaired in mice lacking *Polq*, *Polh*, or both (Martomo et al., 2008). Nevertheless, my research uncovered a role for POLQ in a subset of CSR events as described in Chapter 2 of this thesis (Yousefzadeh et al., 2014).

1.2 POLQ and DNA Repair

Base excision repair

POLQ might serve as a backup DNA polymerase for base excision repair (BER). In mammalian cells, single-nucleotide BER is performed by POLB (pol β) (Prasad et al., 2009; Yoshimura et al., 2006). A fragment of human POLQ was shown to display 5'-deoxyribose-phosphate lyase (dRp-lyase) activity, an activity necessary for single-nucleotide BER (Prasad et al., 2009). This dRp-lyase activity is present within a 24 kDa region of the POLQ polymerase domain (**Figure 2**). This is in contrast to the much stronger dRp-lyase activity of pol β, which is present in a distinct domain separate from the polymerase domain. 5'-dRp lyase activity was still present in a catalytically inactive 98 kDa polymerase fragment of POLQ (Prasad et al., 2009).

In chicken DT40 cells, *Polq* and *Polq Polb* double mutants were examined in regard to BER (Yoshimura et al., 2006). The *Polq* single mutant is sensitive to hydrogen peroxide (which can cause oxidative damage to bases) but not to MMS, although both agents cause damage that is repaired by BER (**Table 2A**). The *Polq Polb* double mutant was more sensitive to MMS than either of the *Polq* or *Polb* single mutants. Extracts from the *Polq* mutant DT40 cell line appear to have a reduced capacity for both single-nucleotide and long patch BER (Yoshimura et al., 2006).

Mouse bone marrow stromal cells lacking *Polq* do not show increased sensitivity to either hydrogen peroxide or paraquat in comparison to controls (Goff et al., 2009). Human SQ20B head and neck cancer cells depleted for POLQ with siRNA are not sensitive to treatment with temozolomide, an alkylating agent that is related to MMS (**Table 2B**) (Higgins et al., 2010b). CH12 mouse B lymphoma cells containing a deletion in an exon required for POLQ catalytic activity are somewhat more sensitive to MMS in the presence or absence of

methoxyamine, which reacts with AP sites and prevents single-nucleotide BER by pol β (Ukai et al., 2006). Taken together, it appears that POLQ is not a major contributor to BER. POLQ may be a backup enzyme for BER in the nematode *C. elegans*, which lacks pol β, although at least one additional DNA polymerase also appears to function in BER in this organism (Asagoshi et al., 2012).

Interstrand crosslink repair

Mus308, the *Drosophila* ortholog of POLQ, is implicated in an alternative form of end joining that may contribute to ICL repair (Chan et al., 2010). *Drosophila mus308*, *C. elegans polq-1*, and *Arabidopsis* TEBICHI are all sensitive to the ICL-inducing agent MMC (**Table 2C**) (Aguirrezabalaga et al., 1995; Inagaki et al., 2006; Muzzini et al., 2008). In vertebrate cells, there are conflicting reports concerning the sensitivity of POLQ to ICL-inducing agents. DT40 cells lacking POLQ are not hypersensitive to cisplatin, a clinically used ICL agent (Kohzaki et al., 2010; Yoshimura et al., 2006). However, deletion of *Polq* slightly sensitized murine CH12B-lymphoma cells to cisplatin and MMC (Li et al., 2011). The *chaos1* mutant mice and cells derived from them exhibited no hypersensitivity to MMC, suggesting that POLQ does not play a major role in the repair of ICLs in mammalian cells (Shima et al., 2004). The investigation of a role for POLQ in the repair of ICL-induced DNA damage is further explored in Chapter 2 (Yousefzadeh et al., 2014).

A

	Vertebrate POLQ knockouts and mutants				
Agent	Chaos1	BMSC	CH12 Pol	CH12 TKO	DT40
IR	+	++	+	+	-
Bleomycin	ND	++	ND	ND	ND
Etoposide	ND	ND	+	+	ND
Peroxide	ND	-	ND	ND	+
Paraquat	ND	-	ND	ND	ND
MMS	ND	ND	-	-	-
MMC	-	ND	ND	ND	-
Cisplatin	ND	ND	+/-	+/-	-
UVC	ND	+	+	ND	-

B

	POLQ-depleted human cells			
Agent	siPOC	siMRC5	SiQ20B	siHeLa
IR	-	-	+	++
Temozolomide	ND	ND	-	ND

C

POLQ knockouts in *Drosophila, C. elegans, and Arabidopsis*			
Agent	Mus308	polq-1	TEBICHI
IR	-	+/-	ND
MMS	-	+	+
MMC	+	ND	+
Cisplatin	+	+	ND
HN2	+	+	ND
UVC	ND	-	ND

Table 2. Sensitivity of POLQ-deficient cells to DNA damaging agents. This table summarizes the sensitivity of POLQ (**A**) vertebrate knockouts and mutants, (**B**) siRNA-depleted cell lines, and (**C**) non-vertebrate animal and plants knockouts to various DNA damaging agents when I initiated my project. Sensitivity of cells to damaging agents were described as (+) hypersensitive, (–) no sensitivity, or (ND) not described.

Tolerance of DNA double strand breaks and alternative end joining (altEJ) of breaks

While evidence for multiple POLQ functions has been discovered, the strongest evidence to date, supports a role for POLQ in the defense against radiation-mediated genomic instability. The *chaos1* (chromosome aberration occurring spontaneously 1) mouse was produced in a phenotype-driven screen (Shima et al., 2003). Male mice were periodically dosed with N-ethyl-N-nitrosourea (ENU) to induce germline mutations before mating. Generation 1 male offspring were backcrossed with generation 2 females to create generation 3 offspring that were potentially homozygous. The frequency of micronuclei (MN) was measured in male generation 3 mice. MN arise from acentric chromosome fragments arising from chromosome breakage or from whole chromosomes that was not properly segregated or packaged into the nucleus at the time of cell division (Nusse et al., 1996). DNA damaging agents induce MN, and those that form shortly before the nucleus is ejected can be assayed in peripheral blood reticulocytes by flow cytometry (Dertinger et al., 1996). The *chaos1* mice show increased spontaneous MN in blood samples, and the level is further elevated by ionizing radiation (IR) and MMC (Shima et al., 2003). The *chaos1* recessive mutation was mapped to a region of chromosome 16 that included the *Polq* gene, and was determined to be a T to C point mutation in exon 19 of mouse *Polq*. This missense mutation converts Ser to Pro at amino acid 1932, located near the C-terminal end of the central domain, adjacent to the polymerase domain of POLQ (**Figure 2A**). Expression of wild-type *Polq* from a bacterial artificial chromosome was able to suppress the *chaos1* phenotype, suggesting that *chaos1* mice possess a mutant allele of *Polq* (Shima et al., 2003).

A $Polq^{-/-}$ knockout mouse was then generated by Naoko Shima in John Schimenti's laboratory, introducing a point mutation in exon 1 to create a premature stop codon and replacing exons 2–5 with a neomycin resistance cassette (Shima et al., 2004). A lack of

complementation was observed between the *chaos1* and *Polq*-null alleles demonstrating that *chaos1* is a mutant allele of *Polq*. The authors also noted a marked developmental disadvantage for *Atm*$^{-/-}$ *Polq*$^{-/-}$double mutant mice. The few mice that did survive displayed signs of severe DNA damage stress, including increased chromosomal instability, decreased body weight and a delayed onset of thymic lymphomas in comparison to *Atm*$^{-/-}$ mice (Shima et al., 2004). Goff *et al.* found that *Polq*$^{-/-}$ mice have increased levels of IR-induced MN and confirmed that the levels of spontaneous MN formation were nearly identical to those reported for *chaos1* mice (Goff et al., 2009; Shima et al., 2004). Further Goff *et al.* found that bone marrow stromal cells from the *Polq*$^{-/-}$ mice were hypersensitive to gamma irradiation and to bleomycin, a drug that causes DNA double-strand breaks (**Table 2A**) (Goff et al., 2009). Inhibition of ATM activation with the drug KU55933 increased radiosensitivity in wild-type but not in *Polq*$^{-/-}$ cells. Li *et al.* constructed a *Polq*-null mutant in mouse CH12F3 B cells. These cells are more sensitive than normal cells to gamma-irradiation and etoposide (Li et al., 2011). POLQ was also uncovered in a siRNA screen of genes that affect the IR sensitivity of human cells. Depletion of POLQ in SQ20B and HeLa cancer cell lines caused an increase in IR-induced γH2AX foci and sensitized the cells to gamma-irradiation (**Table 2B**) (Higgins et al., 2010b). These observations indicate that POLQ has a role in either the tolerance or repair of DNA double-strand breaks.

The mechanism by which POLQ aids in the defense against IR-mediated genomic instability is under investigation. Studies of *Drosophila Mus308* function point towards a role of this gene in an alternative form of end joining which relies upon annealing of microhomologies in order for DNA end joining to occur. Yu *et al.* investigated the inaccurate repair of DSBs mediated by the *I-SceI* nuclease in *Mus308* mutants and found a decrease in a process termed synthesis-dependent microhomology-mediated end joining (SD-MMEJ) (Yu and McVey, 2010). This repair pathway is independent of Ku70, Lig4, and Rad51 and

distinct from “canonical” non-homologous end-joining (cNHEJ) of DNA breaks. SD-MMEJ is related to, or identical to, pathways defined in other biological systems as alternative end joining (altEJ), or micro-homology-mediated end-joining (MMEJ). Analysis of *mus308* mutants showed a decrease in the use of long microhomologies at the junction site and abrogation of complex T-nucleotide insertion events. Further analysis suggested that this altEJ pathway is occurring in parallel to LIG4-mediated canonical end-joining (Chan et al., 2010). Mus308 is also necessary for the incorporation of additional nucleotides (via an altEJ pathway) that allows for retrohoming of group II introns in flies (White and Lambowitz, 2012). Worms lacking the functional POLQ homolog also show a defect in this altEJ repair pathway and exhibit enhanced genomic instability (Koole et al., 2014; Roerink et al., 2014).

One would hypothesize that if cells lacking or with reduced POLQ levels were hypersensitive to a particular DNA damaging agent, then POLQ may play a role in the repair of that specific type of DNA damage. However, although such sensitivity assays have been performed in a number of species (**Table 2**) there is no unified evolutionarily consensus of what genotoxic agents that POLQ mutants are hypersensitive to and what DNA repair pathways it may participate in. In *Drosophila*, Mus308 does not appear to participate in homology-directed repair. *Mus308* mutant larvae are not more sensitive to IR than wild-type, but they are synergistically sensitized to radiation by addition of *spn-A* (Rad51) mutation, indicating that Mus308-dependent repair is independent of homologous recombination in *Drosophila* (Chan et al., 2010). This is in contrast to the *Polq* mutant in chicken DT40 cells, which are only hypersensitive to peroxide and showed decreased levels of HR repair of *I-Sce*I-mediated DNA DSBs (Kohzaki et al., 2010; Yoshimura et al., 2006). Recently, a mutant of POLQ (Z16) in the green algae *C. reinhardtii* was shown to be hypersensitive to the DSB-inducing agent zeocin (Plecenikova et al., 2014). Human and mouse POLQ mutants are radiosensitive, suggesting some divergence of POLQ and Mus308 functions. To understand

a role for POLQ in mammalian cells, hypersensitivity to a broad spectrum of DNA damaging agents needs to be tested in a better controlled manner (See Chapter 2) (Yousefzadeh et al., 2014).

Functions in replication and at the replication fork

There are some suggestions that POLQ might work at the DNA replication fork or in close connection with DNA replication proteins. As described below, *POLQ* gene expression correlates with expression of genes involved in cell cycle control and DNA replication (Pillaire et al., 2009). MRC5VA human lung fibroblast cells were engineered to stably express recombinant POLQ, and this resulted in an increase in γH2AX foci, an increase in chromosomal abnormalities and defective DNA replication fork progression (Lemée et al., 2010). These data show that it is imperative for cells to keep levels of POLQ tightly controlled to avoid detrimental effects on the cell. A recent report provided evidence that POLQ interacts with replication origin factors ORC2 and ORC4 and reduced POLQ levels alter replication timing (Fernandez-Vidal et al., 2014). In the same screen that identified the *chaos1* allele of POLQ, a second mutant termed *chaos3* showed a similar phenotype of increased micronuclei. *Chaos3* has been identified as a hypomorphic allele of *Mcm4*, a component of the helicase operating during semiconservative replication (Shima et al., 2007). In *Arabidopsis*, *TEBICHI* gene mutants show striking defects in G2/M checkpoints that could be due to abnormal DNA replication (Inagaki et al., 2006).

POLQ expression in normal and cancerous tissue

The *POLQ* gene is ubiquitously expressed in human and mouse tissues and cell lines, as measured by northern blot analysis, reverse transcriptase polymerase chain reaction (RT-PCR), and microarrays. Levels of expression appear to be highest in the testis, human placental tissue and hematopoietic cells (Kawamura et al., 2004; Seki et al., 2003;

Shima et al., 2003), but *POLQ* transcripts can be found throughout the body (**Table 3**). POLQ is a member of a cluster of genes in which higher levels of expression correlate with lower patient survival for several cancers. This cluster includes genes involved in DNA replication, repair, recombination and cell cycle control. Higher expression of *POLQ* mRNA was observed in some tumor samples from stomach, lung, and particularly colon cancers compared to matched control non-tumor tissue (Kawamura et al., 2004). Stratification of colon cancers based on *POLQ* expression alone indicated that the "*POLQ* high" group had poorer survival than the "*POLQ* low" group by an average of ~24 months. Another study of colorectal cancer patients indicated that tumors with higher expression of a group of 47 DNA replication-related genes (including *POLQ*) were significantly correlated with poorer patient survival (Pillaire et al., 2009). *POLQ* expression levels alone were correlated with poor patient survival (Pillaire et al., 2009). In an analysis focusing on the expression of DNA polymerases in human breast cancers, Lemée *et al.* found that of all 14 nuclear DNA polymerases, only *POLQ* expression was significantly higher in cancer samples than in normal tissue samples. Stratification, according to expression, showed that higher levels of *POLQ* expression were correlated with poorer survival outcomes (Lemée et al., 2010). Another study of *POLQ* expression in early breast cancer indicated that samples with the highest *POLQ* expression level were in ER negative and high-grade tumors. Further, higher *POLQ* expression was correlated with poor relapse-free survival (Higgins et al., 2010a). Considering that radiation therapy is a frontline treatment for breast cancer and that human cancer cells depleted for POLQ are sensitive to ionizing radiation, POLQ is a molecule of interest in the study of breast cancers and is a potentially druggable target.

In some non-tumor cells, reduced POLQ levels resulting from siRNA treatment had little effect on radiation sensitivity (Higgins et al., 2010b). It would be clinically beneficial if tumor cells but not normal cells could be sensitized by POLQ suppression or inhibition, but it

remains to be seen whether this is generally the case. As described above, genetic ablation of POLQ from normal mouse cell lines leads to radiation hypersensitivity. Despite this radiation sensitivity, the hematopoietic system of *Polq*$^{-/-}$ mice seems resilient. Mice irradiated with a sub-lethal dose of 7 Gy of γ-irradiation showed a subsequent depression of counts for white blood cells, red blood cells, lymphocytes and platelets, as expected after immunoablation. Recovery to normal levels took place in *Polq*$^{-/-}$ mice as well as wild-type mice (Goff et al., 2009). Due to its unique properties and clinical interest, further work to understand POLQ and its role in the cell is pivotal.

Body Site	TPM	Gene EST		EST Pool	Body Site	TPM	Gene EST		EST Pool
adipose tissue	0	0	/	12866	nerve	0	0	/	15535
adrenal gland	30	1	/	32940	ovary	19	2	/	101488
ascites	0	0	/	39834	pancreas	0	0	/	213440
bladder	0	0	/	29860	parathyroid	0	0	/	20594
blood	32	4	/	122252	pharynx	24	1	/	40725
bone	13	1	/	71618	pituitary gland	0	0	/	16526
bone marrow	0	0	/	48737	placenta	0	0	/	283019
brain	1	2	/	1092688	prostate	0	0	/	189536
cervix	0	0	/	48486	salivary gland	0	0	/	20265
connective tissue	6	1	/	149072	skin	4	1	/	210759
ear	124	2	/	16100	spleen	0	0	/	53397
embryonic tissue	14	3	/	212896	stomach	10	1	/	95679
esophagus	0	0	/	20154	testis	9	4	/	435204
eye	14	3	/	208840	thymus	25	2	/	79697
heart	0	0	/	89524	thyroid	0	0	/	46583
intestine	21	5	/	231981	tonsil	0	0	/	17021
kidney	14	3	/	210778	trachea	38	2	/	51780
larynx	1022	24	/	23466	umbilical cord	0	0	/	13764
liver	9	2	/	205291	uterus	30	7	/	232093
lung	5	2	/	334815	vascular	19	1	/	51649
lymph	22	1	/	44302					
lymph node	66	6	/	89748					
mammary gland	26	4	/	151230					
mouth	15	1	/	66150					
muscle	9	1	/	106371					

Table 3. *POLQ* transcripts are found throughout the body. Expressed Sequence Tag (EST) data collected from Unigene for *POLQ*. For each body site, transcript per million (TPM), gene-specific (Gene) EST and total EST's for each body site (EST Pool).

1.3 Aims of this thesis

The full-length POLQ cDNA was cloned in the Wood Laboratory in 2003, sparking research into this specialized DNA polymerase (Seki et al., 2003). From the initial discovery POLQ was unique, it was a fairly large enzyme (predicted molecular weight of 290 kDa) and stood out from the other mammalian DNA polymerases because POLQ also contained a helicase (-like) domain. When I started in the Wood Laboratory, a fair amount of information regarding the *in vitro* biochemical properties of recombinant human POLQ was known; however, a cellular function for POLQ was still unclear. From the outset of the project, Dr. Wood and I felt a profound sense of excitement to determine what mammalian POLQ was doing in the cell. With a minimal amount of knowledge of the DNA repair field, a lot of curiosity, and the steady guidance of a mentor, I embarked on a roughly six year intellectual journey to determine the function of mammalian POLQ.

An early hypothesis that I had in the infancy of my dissertation project was that POLQ played a role in at least one or more DNA repair pathways.

Specific Aim 1 – To determine what DNA repair pathway(s) POLQ participates in. Investigation of this aim was accomplished by analyzing isogenic sets of long-term bone marrow stromal cells derived from *Polq*-null mice and their wild-type littermate controls for hypersensitivity to different DNA damaging agents by clonogenic survival assays. Further mechanistic investigation into the repair of DNA damage utilized *in vivo* analysis of wild-type and *Polq*-null mice, *in vitro* and biochemical assays.

Specific Aim 2 – To determine the biological consequences of the loss of POLQ. *Polq*-deficient mice and cells derived from them were analyzed for hypersensitivity to DNA damage, markers of DNA damage signaling, cytologic markers of chromosomal instability,

and *in vivo* chromosomal translocation. The working hypothesis is that loss of POLQ contributes to impaired DNA repair and enhanced genomic instability.

Specific Aim 3 – To determine the role of POLQ in immunological diversity mechanisms through high-resolution analysis.

In order to investigate the role of POLQ in immunological diversification mechanisms, Peyer's patches and stimulated naïve B cells from wild-type and *Polq*-null mice were analyzed for V region mutagenesis, and class switch recombinant junction composition, respectively, through a mix of standard and next-generation DNA sequencing.

Specific Aim 4 – To determine the interplay between POLN and POLQ.

A colony of *Polq*-null mice were established at Science Park and then crossed to *Poln* knockout mice developed by Dr. Kei-ichi Takata to generate $Poln^{-/-}$ $Polq^{-/-}$ mice. Functional analysis of single mutant mouse controls as well as $Poln^{-/-}$ $Polq^{-/-}$ mice, and cells derived from them, were performed to determine whether POLN and POLQ operate within the DNA repair pathway.

Chapter 2. Mechanism of suppression of chromosomal instability by DNA polymerase POLQ

Note this chapter is based on: Yousefzadeh, M.J., Wyatt, D.W., Takata, K., Mu, Y., Hensley, S.C., Tomida, J., Bylund, G.O., Doublie, S., Johansson, E., Ramsden, D.A.*, et al.* (2014). Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. PLoS Genet *10*, e1004654. With permission from the copyright holder.

2.1 Introduction

DNA double-strand breaks (DSBs) can be formed in cellular genomes by environmental agents such as ionizing radiation. DSBs also arise during normal cellular duplication cycles, when DNA replication stalls at naturally occurring structures or at sites of internally-generated DNA damage. In diversification steps of the mammalian immune system, DSBs are deliberately formed by regulated enzymatic action, to initiate rearrangement of antibody and receptor segments, and as a means to introduce local variation. Because DSBs are toxic and/or mutagenic if not repaired, organisms have multiple mechanisms for DSB repair (Kass and Jasin, 2010; Rassool and Tomkinson, 2010). The primary repair strategies are end-joining mechanisms, which process and rejoin the ends of a DSB, and homologous recombination (HR) pathways which employ an undamaged copy of the DNA as a repair template (**Figure 3**) (Thompson, 2012). End-joining pathways appear to be the first line of defense against DSBs. The most studied pathway is “classical” non-homologous end-joining (cNHEJ), which relies on the DNA-binding Ku70 (*XRCC6*) and Ku80 (*XRCC5*) gene products, and the DNA protein kinase (DNA-PK, *PRKDC*). One or more “alternative” end-joining pathways (altEJ) also exist, which are independent of these factors **(Figure 3**) (Decottignies, 2013; Ramsden and Asagoshi, 2012).

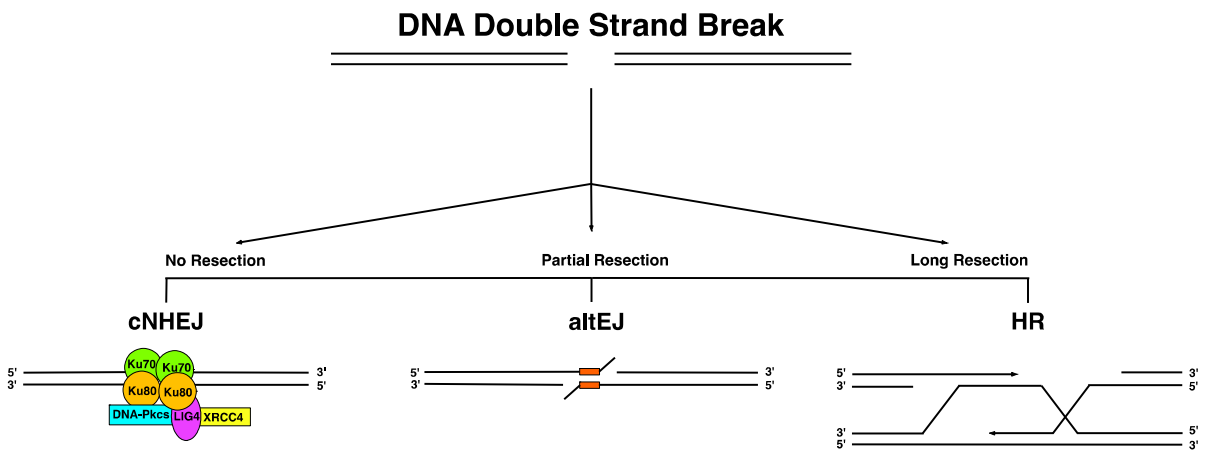


Figure 3. DNA double-strand break repair pathways. Upon recognition of a DNA double-strand break (DSB), DNA damage signaling is activated. Dependent upon cell cycle and sister chromatid availability, the homologous recombination (HR) pathway may be available. Classical non-homologous end joining (cNHEJ), a repair process that is reliant upon the Ku heterodimer and DNA Ligase IV (LIG4), is the predominant repair pathway in mammals and does not rely upon DNA nucleolytic processing to create resected 3' DNA ends. In the absence of, or when either Ku or LIG4 are impaired, alternative end joining (altEJ) anneals minimally resected DNA ends to anneal at existing or newly created microhomologies (5 or nucleotides) as noted in orange. Longer resection is required in order to facilitate HR.

During immunoglobulin diversification, the regional end-joining process of class switch recombination (CSR) replaces one constant region coding exon for another (**Figure 4**). This CSR process is known to occur through both cNHEJ and alternative end joining pathways (Boboila et al., 2012a). In mammalian cells, an alternative end-joining repair pathway repair of DSBs is thought to play a role in driving the formation of chromosomal

translocations, although the specific enzymology is unclear (Simsek and Jasin, 2010; Zhang and Jasin, 2011).

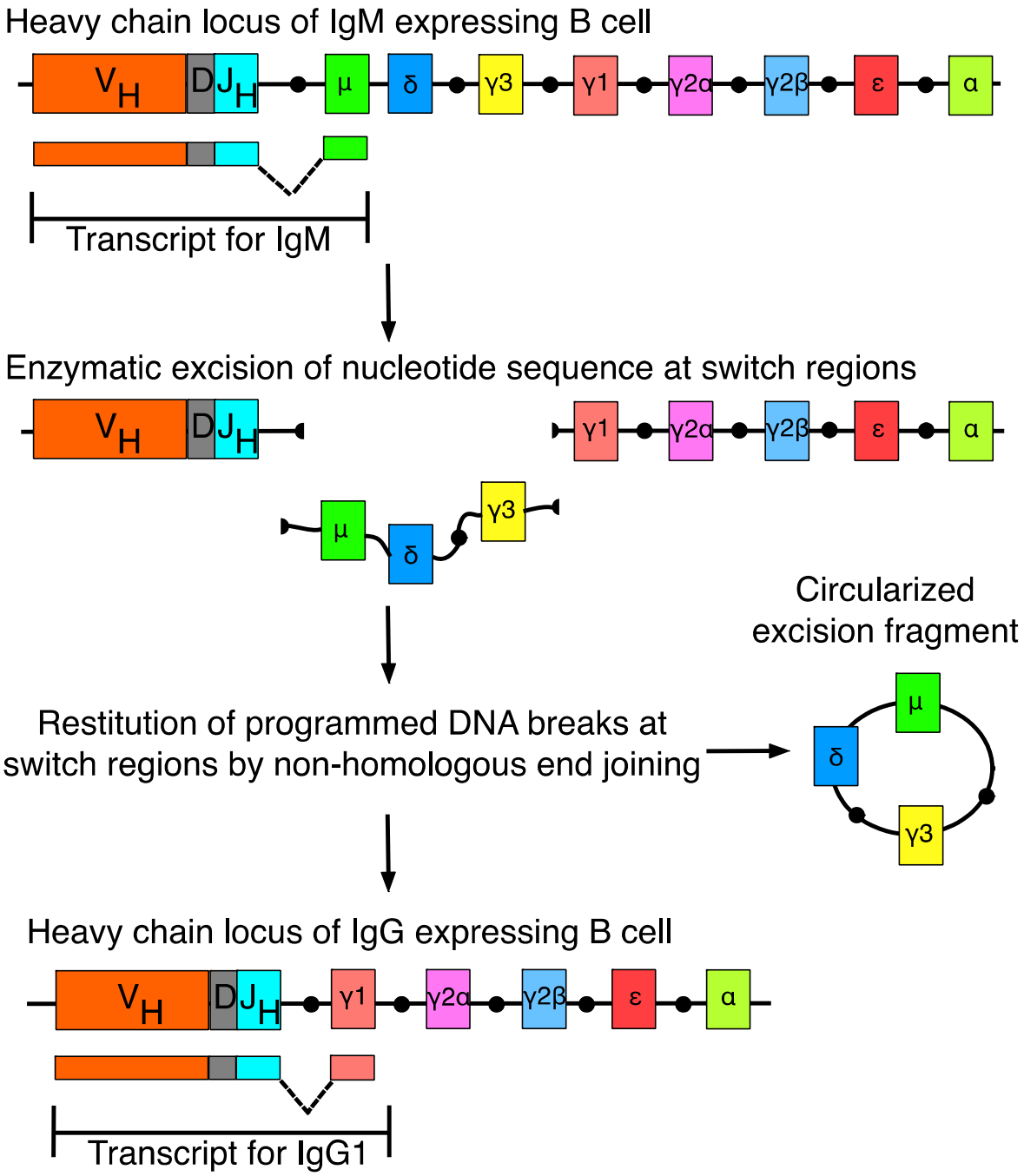


Figure 4. Immunoglobulin class switch recombination. The mammalian immunoglobulin heavy-chain locus of an IgM-expressing B cell is displayed. The heavy-chain variable (V_H), diversity region (D), and joining region (J_H) gene segments are shown in orange, grey, and teal, respectively. The other shaded rectangles represent different constant region exons, and the black dots denote switch regions. After V(D)J recombination, which occurs in the bone marrow, class switch recombination (CSR) takes place in the peripheral lymphoid tissues. Upon AID-induced DNA double-strand breaks at different switch regions, CSR facilitates the joining of a proximal variable region exon to a downstream constant region exon. The diagram denotes the joining of a break between the Sμ and Sγ1 switch region to yield a gene product that would produce functional IgG1. The deleted gene segment is released as a circularized DNA product.

In this chapter, I report experiments that define a specific function and mechanism of action for POLQ in a pathway for alternative end-joining of DNA double-strand breaks in mammalian cells.

2.2 Material and Methods

Cellular proliferation assay

Polq$^{+/+}$ and *Polq*$^{-/-}$ bone marrow stromal cells and mouse embryonic fibroblasts were plated in triplicate (2 ×10^5 cells per 10 cm dish) with 15 mL of complete media (Dulbecco's Modified Eagle Medium+Glutamax, 10% FBS, 1% PenStrep). On the indicated days, cells were trypsinized and live cells were counted using trypan blue exclusion (Countess automated cell counter, Life Technologies). Experiments were repeated three times in order to generate standard deviations. Viability was consistently high for all cell lines examined (>95% trypan blue-excluding cells).

Clonogenic assays with bone marrow stromal cells

For X-irradiation 5 ×10^5 cells were plated on a 10 cm plastic culture dish, and exposed the following day at 2 Gy/min, 160 kV peak energy (Rad Source 2000 irradiator, Suwanee, GA). Cells were then trypsinized for replating. For UVC-irradiation (254 nm peak germicidal lamp) cells were irradiated in 500 µl PBSA (10^5 cells/ml) at 5 J m^{-2} min^{-1} and then plated. For psoralen-UVA treatment, 5 ×10^5 cells were plated on a 10 cm dish and incubated in medium with the indicated concentration of HMT-psoralen for 1 h, the dish was irradiated with 0.9 kJ m^{-2} UVA (365 nm peak, 30 min, 0.5 J m^{-2} sec^{-1}), the psoralen-containing medium was removed, and fresh medium was added before UVA-irradiating for an additional 30 min before replating. Chemicals were added at the indicated concentrations to dishes at the beginning of the experiment. Drugs were solubilized in ethanol (mitomycin c), DMSO (ICRF-193, etoposide, camptothecin, HMT-psoralen, temozolomide, olaparib), or 150 mM NaCl (cisplatin). All chemicals were from Sigma (St. Louis, MO) except ICRF-193 (Enzo LifeScience, Farmingdale, NY), olaparib (AZD2281, Selleck Chemicals, Houston, TX), and mitomycin c (Calbiochem, Darmstadt, Germany). Cells were plated in triplicate in 10 cm dishes and grown for 7–10 days before being fixed and stained with crystal violet. Colonies

of 50 or more cells were quantified and experiments were repeated three times to generate standard deviations. A clonogenic assay was performed with *Rad51D*$^{+/+}$ and *Rad51D*$^{-/-}$ Chinese hamster ovary (CHO) cell lines (generously provided by Dr. Rodney Nairn) exposed to varying concentrations of olaparib.

Micronucleus assay

BMSCs were plated at 1.5 ×10^4 cells per well in chambered slides and treated with the indicated amount of x-rays or etoposide the following day. 48 h later, cells were fixed with 2% para-formaldehyde, stained with DAPI and coverslips were added. Micronuclei were scored by DAPI immunofluorescence. 300 cells were counted per group. Experiments were repeated three times to generate standard deviations.

Human cell transfections

293T cells (kindly provided by Dr. Christopher Bakkenist, University of Pittsburgh Medical School) were plated at 1.5 ×10^5 cells per well in six-well plates and transfected the following day with 2.5 µg of either pCDH (System Biosciences, Mountain View, CA) containing empty vector control, POLQ, POLQ-K121M, POLQ-D2330A,Y2331A, POLQ-S1977P, or POLQ-DM cDNA using Lipofectamine 2000 (Life Technologies) according to manufacturer's specifications. 48 hr after transfection, cells were harvested for RNA isolation (RNeasy, Valencia, CA) or for immunoblotting.

Immunoblotting

For immunoblots, cells transfected in six-well dishes were resuspended in 200 µL of 2× SDS loading buffer (4% SDS, 0.2% bromophenol blue, 20% glycerol, 100 mM Tris HCl pH 6.8, 12% 2-mercaptoethanol) and heated at 95°C for 5 min. 20 µL of extract was separated on a 4–20% polyacrylamide gel, transferred to PVDF membrane, blocked, and blotted with anti-

alpha-Tubulin (Abcam, Cambridge, UK ab4074, 1:10,000), anti-FLAG (Sigma F7425, 1:5,000), anti-PCNA (Santa Cruz, Santa Cruz, CA, sc-56, 1:1,000), anti-HA (RW, 1:10,000), or anti-POLQ (MDACC POLQ20, 1:250) antibodies and corresponding secondary antibodies (Sigma A0168, A0545; 1:10,000) and visualized with ECL reagent (Pierce, Rockford, IL).

Generation and complementation of Polq MEFs

Polq-heterozygous (*Polq*$^{+/-}$) mice (Shima et al., 2004) were obtained from Jackson Laboratories and maintained in a C57BL/6J background. *Polq*-heterozygous mice were crossed to generate *Polq*-null mice and maintained in a holding room that was positive for mouse parvovirus, minute virus of the mouse, and *H. pylori*. Isogenic primary MEFs were generated from 13.5 day pregnant females and cultured in a 2% O_2 atmosphere. MEFs were then transfected with 1 µg of pSV-Tag (Lange et al., 2012; Sobol et al., 1996) and grown in atmospheric oxygen for six population doublings to allow for immortalization. To generate lentivirus used for transduction, 293T cells were cotransfected with psPAX2 (6 µg), pMD2G (6 µg), and pCDH (12 µg) expression vector (See below for construction of expression vectors) using Lipofectamine 2000. One day prior to transduction *Polq*$^{-/-}$ MEFs were seeded into a 10 cm dish at 1.5 ×10^5 cells with 12 mL complete media. 48 h post-transfection virus-containing media was harvested, filtered through a 0.45 µm syringe filter and used to replace the media on the plated MEFs. MEFs were incubated in the virus-containing media for 24 h before being split into T-75 flasks and allowed to grow to 80% confluence before undergoing three weeks of puromycin selection (2.5 µg/mL). Following selection, pure clones were isolated and cultured with complete media containing puromycin (1 µg/mL).

Quantitative real-time PCR analysis of complemented MEFs

RNA isolated from the complemented MEF lines were analyzed for quality and purity using an RNA 6000 Nano kit (Agilent Technologies, Santa Clara, CA). 1 µg of total RNA was used

to generate cDNA using the High Capacity cDNA RT kit (Life Technologies). qPCR analysis was performed in triplicate using the ABI Prism 7900 HT thermocycler and the following Taqman Probe set or primer set with iTAQ SYBR Green Supermix with ROX (Bio-Rad, Hercules, CA): MmPolQ_FWD 5′-GGCTCTGAAGAACTCTTTGCCTTT-3′, MmPolQ_REV 5′-GCTGCTTCCTCTTCTTCATCCA-3′, probe 5′-TCCGGGCACTTTTG-3′; HsPOLD1_FWD 5′-CGACCTTCCGTACCTCATCTCT-3′ HsPOLD1_R 5′-ACACGGCCCAGGAAAGG-3′, probe 5′-CCCTCAAGGTACAAACAT-3′; Qexon FWD 5′-TGCCTTTCAAAAGTGCCCGGAAGGC3′, Qexon REV 5′-TGCCAGTCACCCANATAGTTCNCAT-3′ (where dSpacer is represented as N to compensate for mismatches between the mouse and human sequences). Data were analyzed using the ΔΔCt method. For absolute quantification, titration of pCR-XL-TOPO/MmPolQ and pET/MmPold1 plasmids were used to generate standard curves for expression. Transcript abundance was determined by extrapolation from linear regression analysis of best-fit lines from titration experiments. *GAPDH* was used as an internal control in all experiments.

Bleomycin sensitivity

Complemented MEF lines were plated in triplicate into white 96 well plates at 1250 cells per well and grown overnight using complete media containing puromycin (1 μg/mL). The following day, cells were cultured with complete media containing the indicated amounts of bleomycin (dissolved in 150 mM NaCl) for 24 hr before the media was replaced. Cells were allowed to recover for 72 hr before cellular viability was measured using the ATPlite 1Step kit (Perkin Elmer, Waltham, MA) using a Biotek plate reader. Experiments were repeated three times.

Immunofluorescence

Complemented MEF lines were plated at a density of 1.5 $\times 10^4$ cells per well in 4-well chamber slides and the following day were irradiated with either 0 or 6 Gy of x-rays. Media was changed and cells were allowed to recover for 48 h after damage before fixation with 2% para-formaldehyde and permeabilization with Triton X-100. Samples were blocked with donkey serum for 30 minutes before being incubated overnight with primary antibodies against 53BP1 (Bethyl, Montgomery, TX, A300-272A, 1:500) and γH2AX (EMD Millipore 05-636, 1:400). Cells were later incubated with AlexaFluor-488 goat-anti-mouse or AlexaFluor-594 goat-anti-rabbit secondary (Life Technologies, 1:1000) and then stained with DAPI before coverslips were placed. Cells were scored for DSBs by enumerating the percentage of cells with >2 53BP1 foci and >2 γH2AX foci (Lange et al., 2012; Ward et al., 2003). The majority of cells that contained >2 foci for each of the DSB markers, exhibited colocalization of the foci. Cells with pan-staining of γH2AX were not included in the analysis as they are proposed to represent pre-apoptotic cells (Marti et al., 2006). Many of the cells with 53BP1 foci, exhibited enlarged foci that are associated with nuclear OPT (Oct-1, PTF, transcription) domains that sequester damaged DNA in G1 (Harrigan et al., 2011; Lukas et al., 2011). Thus, most of the MEFs that were foci positive contained DSBs (Harrigan et al., 2011). DAPI-stained micronuclei were also scored. For each experiment 250 cells were scored for three independent experiments for a total of 750 cells.

DNA polymerase assays

Purified POLQ was as described (Hogg et al., 2011). Klenow Fragment (3′→5′ exo-) was purchased from NEB. POLQ was diluted in buffer containing 30 mM Tris-HCl pH 8.0, 50 mM NaCl, 2 mM DTT, 10% glycerol, 0.01% Triton X-100, and 0.1% BSA. Klenow Fragment (3′→5′ exo-) was diluted in buffer containing 25 mM Tris-HCl pH 7.4, 1 mM DTT, and 0.1 mM EDTA. POLQ reaction mixtures (10 µl) contained 20 mM Tris-HCl pH 8.8, 4% glycerol,

2 mM dithiothreitol (DTT), 80 µg/ml bovine serum albumin (BSA), 8 mM $MgCl_2$, 0.1 mM EDTA, 100 µM of each dNTP, 30 nM of the primer-template or primer (see below). Klenow Fragment (3′→5′ exo-) reaction mixtures (10 µl) contained 10 mM Tris-HCl pH 7.9, 50 mM NaCl, 1 mM DTT, 10 mM $MgCl_2$, 100 µM of each dNTP, and 30 nM of the primer-template or primer. After incubation at 37°C for 10 min for a 16-1+PA42 substrate or 20 min for 16-1, C20, and C19THF substrates, reactions were terminated by adding 10 µl of formamide stop buffer (98% formamide, 10 mM EDTA pH 8.0, 0.025% xylene cyanol FF, 0.025% bromophenol blue) and boiling at 95°C for 3 min. Products were electrophoresed on a denaturing 20% polyacrylamide-7 M urea gel, exposed to BioMax MS film, and analyzed with a STORM 860 Phosphor Imager (Molecular Dynamics).

Oligonucleotide Substrates

Note: Substrates designed by Dr. Kei-ichi Takata and MJY. Assays performed by Dr. Kei-ichi Takata and analyzed by Dr. Kei-ichi Takata and MJY. Gel purified primer oligonucleotides 16-1 and PA42 were purchased from Bio-Synthesis, C20 was purchased from Midland, C19+THF was from IDT. 16-1, C20, C19+THF were 5'-labeled using polynucleotide kinase (Promega) and [γ-^{32}P] dATP (PerkinElmer). 5'-[^{32}P]-labeled 16-1 primer was annealed to PA42 to generate a double stranded DNA substrate as described (Takata et al., 2006)

16-1 (primer)+PA42 (template) substrate:

5'-^{32}P-CACTGACTGTATGATG-3'

3'-GTGACTGACATACTACTTCTACGACTGCTC-5'

16-1 substrate (primer only):

5'-^{32}P-CACTGACTGTATGATG-3'

C20 substrate (primer only):

5'-^{32}P-CCCCCCCCCCCCCCCCCCCC-3'

C19+THF (tetrahydrofuran, an AP site analog) substrate (primer only):

5'-^{32}P-CCCCCCCCCCCCCCCCCCC[THF]-3'

Sequencing of DNA extended by POLQ

Note: Substrate and experimental procedures were designed by Dr. Göran Bylund and Dr. Erik Johansson. Polymerase assay and sequencing of extension products were performed by Dr. Göran Bylund and Sara Lundsten. Data was analyzed by Dr. Göran Bylund and Dr. Erik Johansson. A truncated version of human POLQ (residues 1792-2590) was expressed in *E. coli* and purified as described (Hogg et al., 2011). Oligonucleotides were purchased from Eurofins MWG operon (Germany). The 200 µl primer extension reaction contained: [20mM Tris-HCl (pH 8.8), 4% glycerol, 10 mM $MgCl_2$, 0.1 mM EDTA, 80 □g/ml bovine serum albumin (BSA), 10 pmol of the tetrachlorofluorescein (TET) labelled oligonucleotide "no-hairpin" (5'-(TET)-GTTTGTGAGTTTCC-3') and 1.25 □g $POLQ_{(1792-2590)}$]. The primer extension reaction was incubated at 22 °C for 2 h and terminated by the addition of 10 µl 0.5 M EDTA and POLQ was heat inactivated for one min at 100 °C. As a control, 10 µl of the reaction mixture was separated on a polyacrylamide sequencing gel and the amount of POLQ extended products was evaluated with a Typhoon 9400 Variable Mode Imager (GE Healthcare). The remaining DNA was precipitated overnight at -20 °C with 4 µg carrier tRNA, 4 µg glycogen, 0.25 volumes of 5 M ammonium acetate and 2.5 volumes of ethanol. Precipitated DNA was dissolved in 100 µl 10mM Tris-HCl (pH 8.0) and used as template for terminal nucleotidyl transferase (TdT) (Thermo Scientific). 40 µl TdT reactions contained 20 µl POLQ extended oligonucleotide, 1X TdT buffer, 3.2 □M either of dATP/dTTP or 1.5□M either of dCTP/dGTP and 30 U TdT. The reactions were incubated at 37 °C for 15 min followed by heat inactivation of TdT at 70 °C for 10 min. In control experiments TdT reactions were performed with a no-hairpin oligonucleotide, which had not been extended by POLQ, as template. 10 µl from each reaction was separated on a polyacrylamide

sequencing gel and the amount of TdT extended products was visualized and evaluated with a Typhoon 9400 Variable Mode Imager. POLQ + TdT extended no-hairpin DNA and only-TdT extended no-hairpin DNA were amplified by polymerase chain reaction (PCR). 25 µl PCR reactions contained: 0.2 □l from each respective TdT reaction, 5 pmol of the oligonucleotide EcoRI-nohairpin (5'- TTTTGAATTCGTTTGTGAGTTTCC-3'), 5 pmol of either HindIII-25A (5'-TTTTAAGCTTAAAAAAAAAAAAAAAAAAAAAAA-3'), HindIII-25T-R (5'-TTTTAAGCTTTTTTTTTTTTTTTTTTTTTTTTT-3'), HindIII-15C-R (5'-TTTTAAGCTTCCCCCCCCCCCCCC-3') or HindIII-15G-R (5'-TTTTAAGCTTGG-GGGGGGGGGGGGG-3'), 1X GoTaq buffer, 200 µM of each dNTP and 1 U GoTaq DNA polymerase (Promega). PCR cycling conditions were: five cycles (94 ^{o}C 30 s, 50 ^{o}C 30 s and 72 ^{o}C 10 s) followed by 30 cycles (94 ^{o}C 30 s, 55 ^{o}C 30 s and 72 ^{o}C 10 s). PCR reactions were purified with QIAquick Nucleotide Removal Kit (Qiagen, Germany) and digested with *EcoR*I and *Hind*III. Digested PCR products were then separated on a 2% agarose gel containing Gel Green Nucleic Acid Stain. DNA was excised, purified with QIAquick Gel Extraction Kit (Qiagen, Germany), ligated into an *EcoR*I-*Hind*III digested pUC18 vector and transformed into competent XL1-Blue bacterial cells. Plasmid containing bacterial clones were selected on LB plates containing 100µg/ml carbenicillin and plasmids were isolated with the E-Z 96 fastfilter plasmid kit (Omega bio-tek) and sent for sequencing (Eurofins MWG Operon, Germany).

Extrachromosomal substrate assays

Note: Substrates were designed by David Wyatt and Dr. Dale Ramsden, assays were performed by David Wyatt, and data was analyzed by David Wyatt and Dr. Dale Ramsden. *Polq*-null and complemented cell lines were generated and provided by MJY. A dermal fibroblast line from Ku70 and p53 deficient mice (the gift of Dr. P. Hasty, University of Texas Health Sciences Center) was transduced with empty vector (pBABE-puro) retrovirus or a

retrovirus expressing mouse Ku70. Substrates were generated by ligating short linkers to the head and tail of a 556 bp linear double-stranded DNA fragment. Linkers possessed 16–17 bp of double-stranded DNA and either 6 or 45 nt 3′ single-stranded overhangs. The linkers with 6 nt overhangs were made by annealing 5′-AGTCTGAGATGGGTGTGAGATCTGC-3′ to 5′-CACTCTCTCACACCCATCTTA-3′ (“head” linker), and 5′-TGACTATACAGCTAAGCGATGATGCAG-3′ to 5′-CATCGCTTAGCTGTATA-3′ (“tail” linker). The linkers with 45 nucleotide 3′ overhangs were generated by annealing 5′-AGTCTGAGATGGGTGTGAGAGTGAAGATCCTCACCTT CGGAGTACTCCTTCTTTTGAGATCTGC-3′ to 5-CTCACACCCATCTCA-3′ (“head” linker) and5′-TGACTATACAGCTAAGCGATGCTCTCACCGAGCGTAT CTGCTGTGTTGTGGATGAATTAGATGCAG-3′ to 5′-CATCGCTTAGCTGTATA-3′ (“tail” linker). Excess linker was removed by QiaQuick purification and substrate purity validated by polyacrylamide gel electrophoresis. 75 ng of substrate was mixed with 1.1 µg of supercoiled pMAX-GFP (Lonza) plasmid carrier and introduced into 2 $\times 10^5$ cells in a 10 µl volume by electroporation with one 30 ms 1350 V pulse (Neon, Invitrogen). Cells were harvested after incubation for 1 h at 37 °C, washed, resuspended in Hank's buffered saline solution supplemented with 5 mM $MgCl_2$, and extracellular DNA digested by incubation with 6.25 U Benzonase (Novagen) for 15 min at 37 °C. Cells were pelleted and DNA purified with the Qiamp kit (Qiagen). Joining efficiency was determined by quantification of head-to-tail junctions by qPCR using primers that either anneal within double-stranded flanks (5′-CTTACGTTTGATTTCCCTGACTATACAG-3′ and 5′- GCAGGGTAGCCAGTCTGAGATG-3′; 6 nt overhang, Figure 5B) or, for the 45 nt overhang substrate only, which anneal to overhang sequence (5′- TAAGCGATGCTCTCACCGAG and 5′-GATGGGTGTGAGAGTGAAGATC; 45 nt overhang, Figure 5C). Results from electroporated samples were further corrected for differences in transfection and sample processing

efficiency qPCR specific for substrate (5′- GGCACTCTCCAAGGCAAAGA and 5′-ACATGTCTAGCCTATTCCCGGCTT).

B cell culture and CSR analysis

Note: B cell isolation and culture was done with the assistance of Dr. Kevin McBride and Sean Hensley. FACS analysis was performed by Dr. Kevin McBride and DNA isolation, sequencing, and sequence analysis was performed by MJY. B cells were isolated from mouse spleens (n = 6 per genotype) by surgical dissection of mice to isolate spleens. Spleens were then ground up on a 70 µM filter using a syringe plunger to release B cells. Naïve B cells were isolated by anti-CD43 (Ly-48) bead negative selection using a MACS column (Miltenyi Biotec, Dallas, TX) and stimulated for class-switching in culture for 72 h. Where indicated, cultures were incubated with DNA-PKcs inhibitor 20 µM NU7026 (Tocris, Bristol, UK) dissolved in DMSO, or mock-treated. The stimulation procedure and flow-sorting for CSR analysis was as described (Callen et al., 2009; Reina-San-Martin et al., 2003). Prior to this analysis, cells were counted; numbers and viability were similar for all groups. Sµ-Sγ1 CSR junctions were amplified by PCR using the following conditions for 25 cycles at 95 ℃ (30 s), 55 ℃ (30 s), 68 ℃ (180 s) using the primers (FWD 5′-AATGGATACCTCAGTGGTTTTTAATGGTGGGTTTA-3′; REV 5′ CAATTAGCTCCTGCTCTTCTGTGG-3′) and *Pfu* Turbo (Stratagene, La Jolla, CA). To the PCR reaction, 5 U of *Taq* polymerase (Promega, Madison, WI) was added and incubated at 72 ℃ for 10 min. The resulting product was TOPO TA cloned and transformed into Top10 *E. coli* cells (Life Technologies, Carlsbad, CA) and plasmids were purified and sent for sequencing using M13 FWD and REV primers in addition to the amplification primers for sequencing. 100 clones for each group were analyzed for mutations, deletions, insertions, and sequence overlaps at the junction and both 30 nt upstream and downstream of the junction. p-values were determined by using the two-tailed Fisher's exact test.

Translocation assay

Note: B cell isolation and culture and DNA isolation was performed by MJY. Sean Hensley performed PCR amplification of translocation derivative and Dr. Yunxiang Mu performed Southern Blotting. MJY and Dr. Kevin McBride analyzed the data. Naïve B cells from three pairs of *Polq*$^{+/+}$ and *Polq*$^{-/-}$ mice were harvested as above, cultured for 72 h, and DNA was isolated. 32 separate PCR reactions, each containing the genome from 1 $\times 10^5$ cells, was performed with primers to amplify *Myc*/*IgH* translocations and amplified translocations were verified by Southern blotting using internal probes to the *Myc* and *IgH* loci as described (Kovalchuk et al., 1997; Loveday et al., 2011). Three independent experiments were performed and the p-value was determined using the two-tailed Fisher's exact test. %IgG1 was also measured as an internal control to ensure the B cells from each genotype were switching at a comparable level.

Generation of POLQ monoclonal antibody

The QL fragment of POLQ (Hogg et al, 2011) was used to immunize a mouse, and a monoclonal antibody that recognizes POLQ was generated. The epitope recognized by this antibody (MDACC POLQ20) was mapped using peptide arrays to amino acid residues 2147-2151. For more details see Chapter 6.

Generation of POLQ expression constructs

Note: All cloning and sequencing was performed by MJY. Dr. Junya Tomida generously provided the pCDH vector backbone with the C-terminal FLAG-HA epitope tags. The POLQ cDNA was cloned from pFastbacHTc/POLQ (Seki et al., 2003) using KOD Xtreme (EMD Millipore, Billerica, MA) polymerase using the following primers and conditions (FWD 5’ CGGGATCCTAGCACCATGAATCTTCTGCGTCGGAGTG□ 3’, REV 5’-

CACATCAAAGTCCTTTAGCTCTCCCGCCGGCGTTTTCCTTTT-3'); 20 cycles at 95 ℃ (30 s), 60 ℃ (30 s), 68 ℃ (10 min). PCR products were subsequently digested with *BamH*I and *Not*I and ligated into a similarly digested pBluescript SKII+ vector. pBluescript/POLQ was propagated and sequenced to ensure no mutations, before the POLQ cDNA was digested with the aforementioned restriction enzymes and ligated into pCDH-FHC vector. The final vector pCDH-FHC/POLQ was propagated and sequenced for insert verification and mutation analysis. Engineering of the mutant pCDH/POLQ vectors is detailed below. The individual domains of POLQ (Helicase-like, Central, and Pol domain) were amplified from pBluescript/POLQ using KOD Xtreme and the following primers before being digested with *Bam*HI and *Not*I, then ligated into pBluescript:

Helicase-like (HLD) domain

CDH HLD5' CGGGATCCTAGCACCATGAATCTTCTGCGTCGGAGTG;

CDH HLD3' TGCCCATGCGGCCGCGATTATAAAACTCTCGAGCAATGATATCTGC

Central (CEN) domain

CDH CEN5' AGGAAAGGATCCATTGCTCGAGAGTTTTATAATCTACATCATCAAGC; CDH CEN3' TCGTTGCGGCCGCTTCCTTTCGCAAGCAAGATTGAAGGTACCACATC

Polymerase (POL) domain

CDH POL5' AGGAAAGGATCCGATGTGGTACCTTCAATCTTGCTTGCGAAAGGAA;

CDH POL3' AAAAGGAAAagcggccgcCACATCAAAGTCCTTTAGCTCTCC

Vectors were sequenced to verify no mutations before site-directed mutagenesis was performed with the following plasmid-primer combinations using KOD Xtreme:

POLQ-K121M

pBluescript/HLD

Forward: 5' CTACAAGTGCTGGGATGACTCTTGTGGCAG 3'□;

Reverse: 5' CTGCCACAAGAGTCATCCCAGCACTTGTAG 3'□

POLQ-D2330A/Y2331A

pBluescript/Pol

Forward: 5' GTTCAATACTGGCTGCTGCCGCCTCTCAGCTTGAACTGAG 3'□;

Reverse: 5' CTCAGTTCAAGCTGAGAGGCGGCAGCAGCCAGTATTGAAC 3'

POLQ-S1977P

PBluescript/Pol

Forward: 5' CTATAAAATTCTTCTTCTTCCTTGTGGCATCTCCTTGGAG 3';□

Reverse: 5' CTCCAAGGAGATGCCACAAGGAAGAAGAAGAATTTTATAG 3'□

Following mutational verification, the individual domains were digested with the corresponding restriction enzymes before being ligated into pCDH-FHC that was previously digested with *BamH*I and *Not*I:

pBluescript/HLD (*BamH*I, *Xho*I), pBluescript/CEN (*Xho*I, *Kpn*I),

pBluescript/POL (*Kpn*I, *Not*I)

Following ligation, plasmids were sequenced to ensure proper ligation and maintenance of engineered mutations.

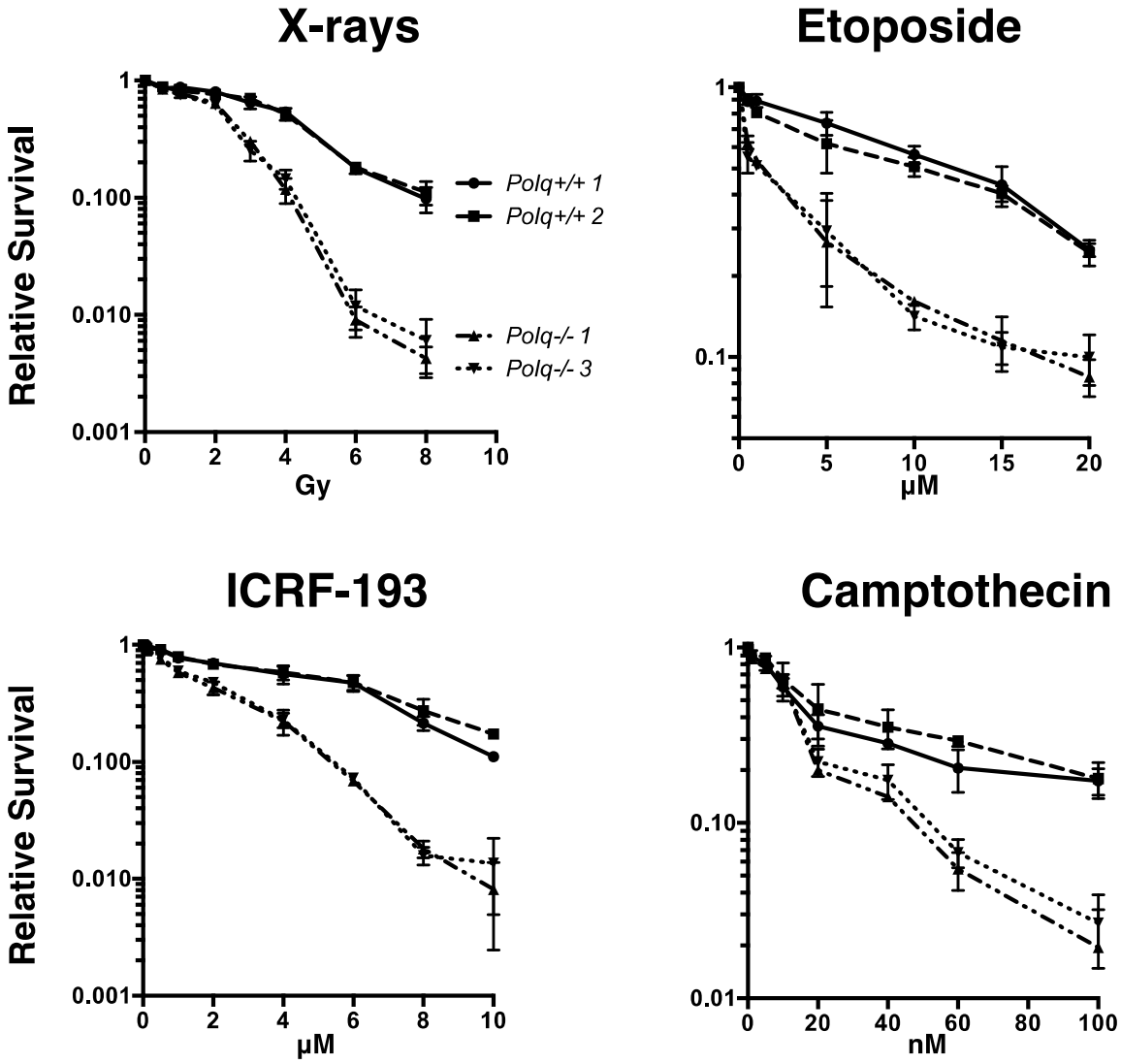
2.3 Results

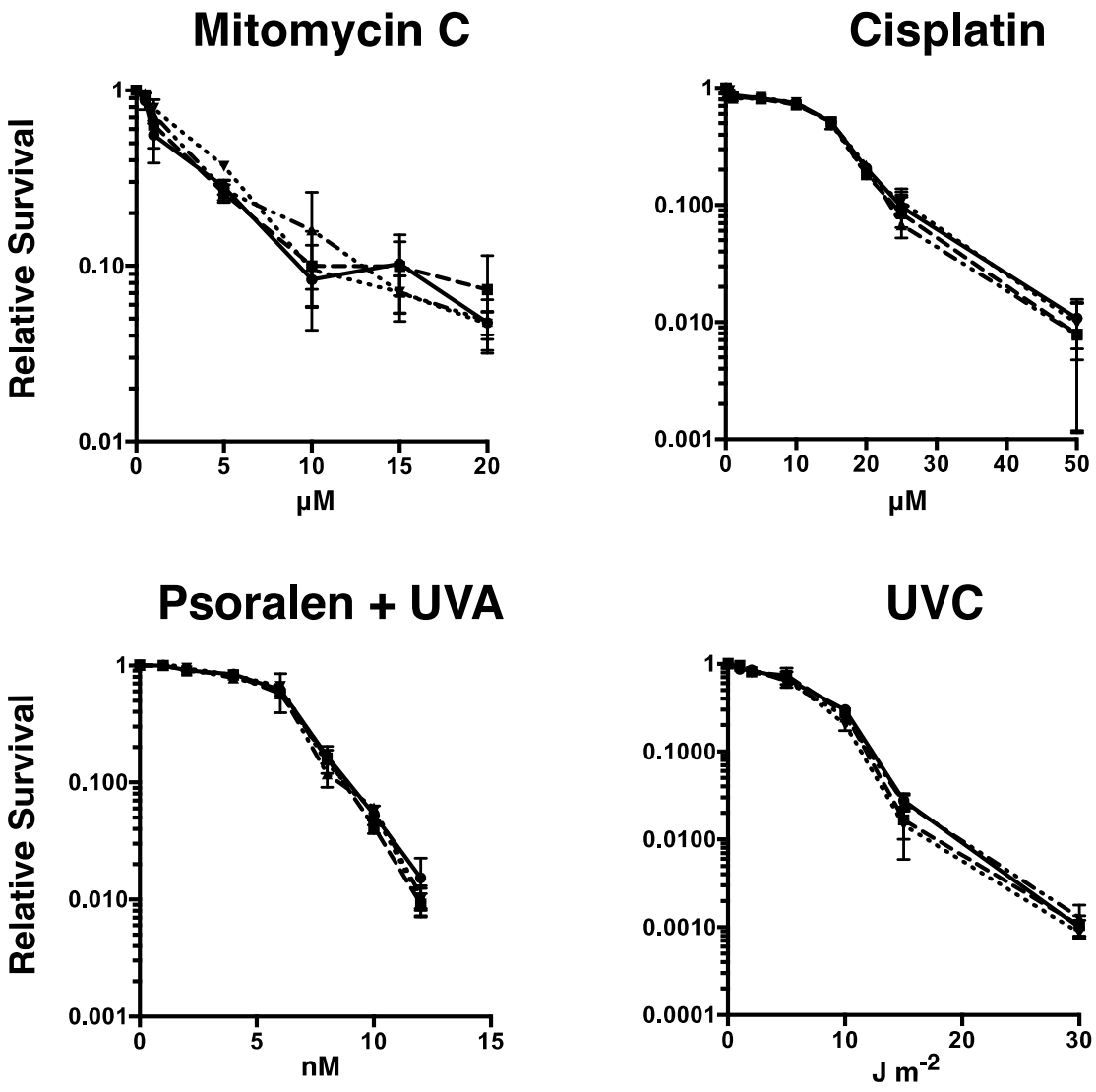
Hypersensitivity to DNA strand-breaking agents in the absence of Polq

As discussed in Chapter 1, mutants of POLQ in different species exhibit a broad range of hypersensitivity to various DNA damaging agents (**Table 2**) (Yousefzadeh and Wood, 2013). To clarify the cellular role of POLQ in response to DNA damage, we measured the sensitivities of *Polq*-null and *Polq*-proficient bone marrow stromal cell (BMSC) lines to various DNA damaging agents. BMSC lines lacking POLQ were previously shown to exhibit hypersensitivity to ionizing radiation (Goff et al., 2009; Higgins et al., 2010b), Interestingly, one of the reports found that only cancer cells with reduced POLQ levels were hypersensitive to IR. "Normal" cell lines suppressed for POLQ by siRNA did not show hypersensitivity to IR under the same conditions (Higgins et al., 2010b). These contrasting results warrant further investigation.

We found that *Polq*$^{-/-}$ cells were also hypersensitive to agents that directly cause DNA breaks (**Figure 5**). These included topoisomerase II inhibitors etoposide and ICRF-193 (Huang et al., 2001), and camptothecin, a topoisomerase I inhibitor. The observation of hypersensitivity of cells to camptothecin in the absence of *Polq* was intriguing. Camptothecin generates single-strand breaks and is only thought to induce DSBs indirectly when the replication or transcription machinery collides with the trapped cleavage complex (Strumberg et al., 2000). DSBs could also form if two trapped Topoisomerase I complexes on different strands are near each other, or if the single-stranded DNA near the trapped cleavage complex falls prey to nucleases. Loss of *Polq* did not cause hypersensitivity to agents that primarily form DNA replication-blocking adducts on one DNA strand including, for example ultraviolet radiation and the methylating agent temozolomide. The *Polq*$^{-/-}$ cells were also not more sensitive than control cells to mitomycin C, cisplatin, and UVA-photoactivated psoralen plus UVA, all of which induce some interstrand DNA crosslinks (ICLs) (**Figure 5**). In contrast to a prior report (Higgins et al., 2010b), our findings show that at least some *Polq*$^{-/-}$ non-

cancer cell lines do, in fact, exhibit hypersensitivity to IR (**Figure 5**) (Yousefzadeh et al., 2014).




Mitomycin C
Cisplatin
Psoralen + UVA
UVC
Relative Survival
μM
nM
J m-2

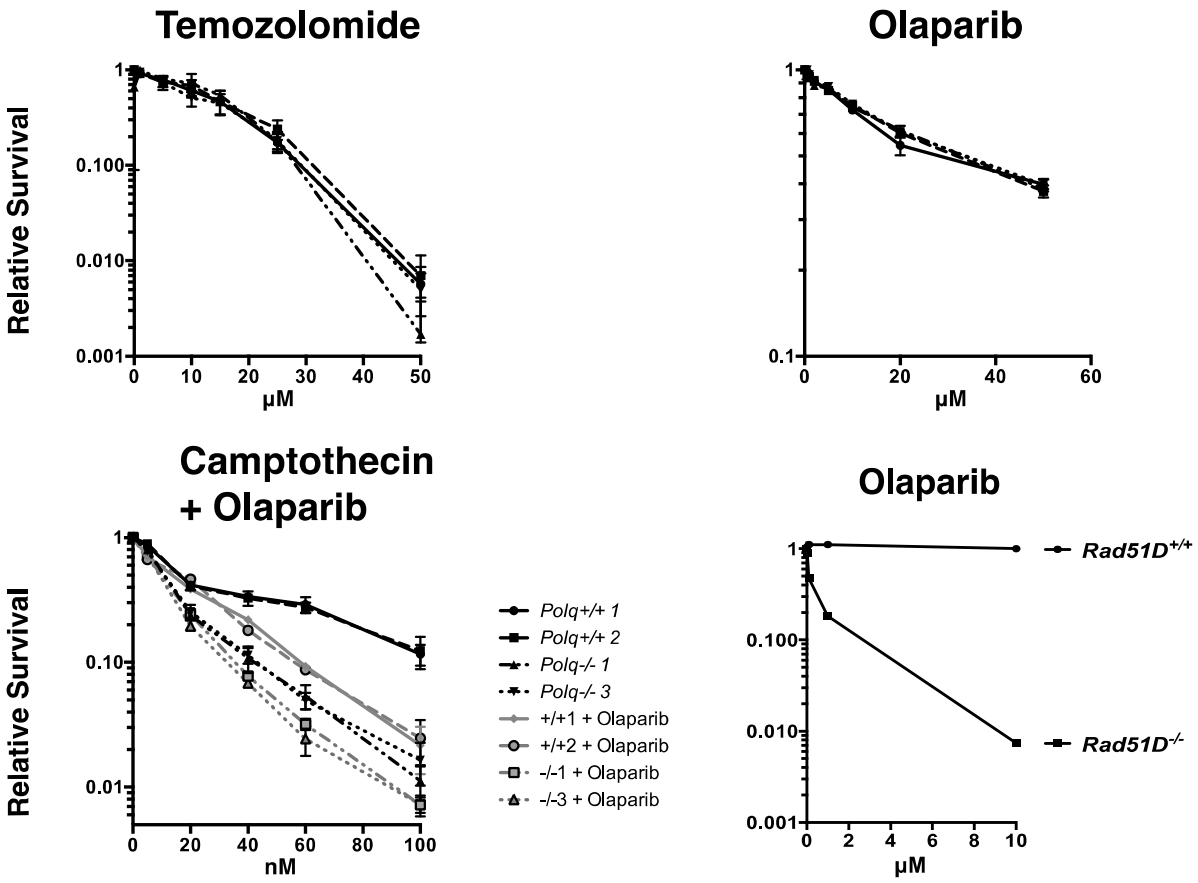


Figure 5. Hypersensitivity of *Polq*$^{-/-}$ bone marrow stromal cells to DNA strand-breaking agents. BMSCs were exposed to x-rays or UVC at the indicated doses, and to etoposide, ICRF-193, camptothecin, olaparib, temozolomide, mitomycin c, cisplatin, and HMT psoralen+UVA at the indicated concentrations and plated in triplicate. Two isogenic bone marrow stromal cell lines were used for each genotype: *Polq*$^{+/+}$ or *Polq*$^{-/-}$. Colonies were stained with crystal violet and counted seven to ten days later. Experiments were repeated three times. Circles, *Polq*$^{+/+}$ clone 1; Squares, *Polq*$^{+/+}$ clone 1; Triangles, *Polq*$^{-/-}$ clone 1; inverted triangles, *Polq*$^{-/-}$ clone 3. Clonogenic assay of *Rad51D*$^{+/+}$ and *Rad51D*$^{-/-}$ CHO cells treated with the indicated doses of olaparib, as a positive control (Loveday et al., 2011). All clonogenic survival assays performed by MJY except those with the *Rad51D* CHO cells, which were performed by Megan Lowery.

These data indicate that POLQ is most important for a process conferring resistance to DNA strand-breaks, particularly double-strand breaks. Sensitivity to olaparib, a PARP inhibitor, was measured because homologous recombination mutants exhibit hypersensitivity to this agent. A current theory is that cells are ill-equipped to deal with the increase conversion of single-strand breaks that are converted to DSBs when PARP is impaired (Bryant et al., 2005; Farmer et al., 2005; Loveday et al., 2011). Cells lacking *Polq* were not hypersensitive to the PARP inhibitor olaparib (**Figure 5**) while control *Rad51D*-defective cells were hypersensitive (**Figure 5**). This suggests that POLQ does not function in the BRCA/homologous recombination pathway (Farmer et al., 2005). POLQ-proficient cells treated with both olaparib and camptothecin were significantly sensitized compared to camptothecin alone. However, addition of olaparib to *Polq*-null cells only modestly increased the sensitivity to camptothecin (**Figure 5**). Besides being involved in single-strand break repair, PARP is also thought to operate within altEJ, that functions independently of cNHEJ factors Ku and DNA Ligase 4. Consequently, PARP and POLQ may operate within the same subpathway of DNA repair.

Loss of Polq enhances chromosomal instability in somatic cells

Previous work revealed an increase in micronuclei in *Polq*-defective cells in peripheral erythrocytes. To determine whether this phenomenon is also true in other cell types, matched wild-type and $Polq^{-/-}$ BMSC lines were exposed to etoposide or x-rays, and the number of cells with micronuclei after 48 h were enumerated (**Figure 6A and B**). *Polq*-null cells exhibited a ~3 fold increase in the frequency of spontaneous micronuclei formation (**Figure 6C**). Upon exposure to DNA damaging agents, the percentage of cells with micronuclei increased about 1.5-fold more per unit of damage for $Polq^{-/-}$ cells in comparison to $Polq^{+/+}$ cells (**Figure 6A and B**). This shows that the susceptibility to micronucleus

formation in *Polq*$^{-/-}$ cells is not confined to cells of the hematopoietic lineage, but occurs also in cultured cells, including fibroblast-like BMSCs.

Cells lacking *Polq* were analyzed for their ability to proliferate in culture. Two independent BMSC lines devoid of *Polq* expression proliferated at a rate comparable to a pair of wild-type control cells. The *Polq*$^{-/-}$ BMSCs showed only a 5% lengthening of population doubling times (**Figure 6D and E**). I extended this analysis to isogenic immortalized mouse embryonic fibroblast (MEF) cell lines (**Figure 6F and G**). Again, *Polq*$^{-/-}$ cells divided at a rate comparable to *Polq*$^{+/+}$ cells. These findings fit with previous observations from the Wood laboratory that hematopoietic cell counts in irradiated *Polq*-null mice recovered at rates comparable to wild-type mice (Goff et al., 2009). I observed no major alterations in growth or development in unchallenged *Polq* null or mutant mice, consistent with previous reports (Shima et al., 2003; Shima et al., 2004; Zan et al., 2005). These observations indicate that despite some increased chromosomal instability, POLQ-defective cells originating from a variety of tissues can proliferate at near-normal rates.

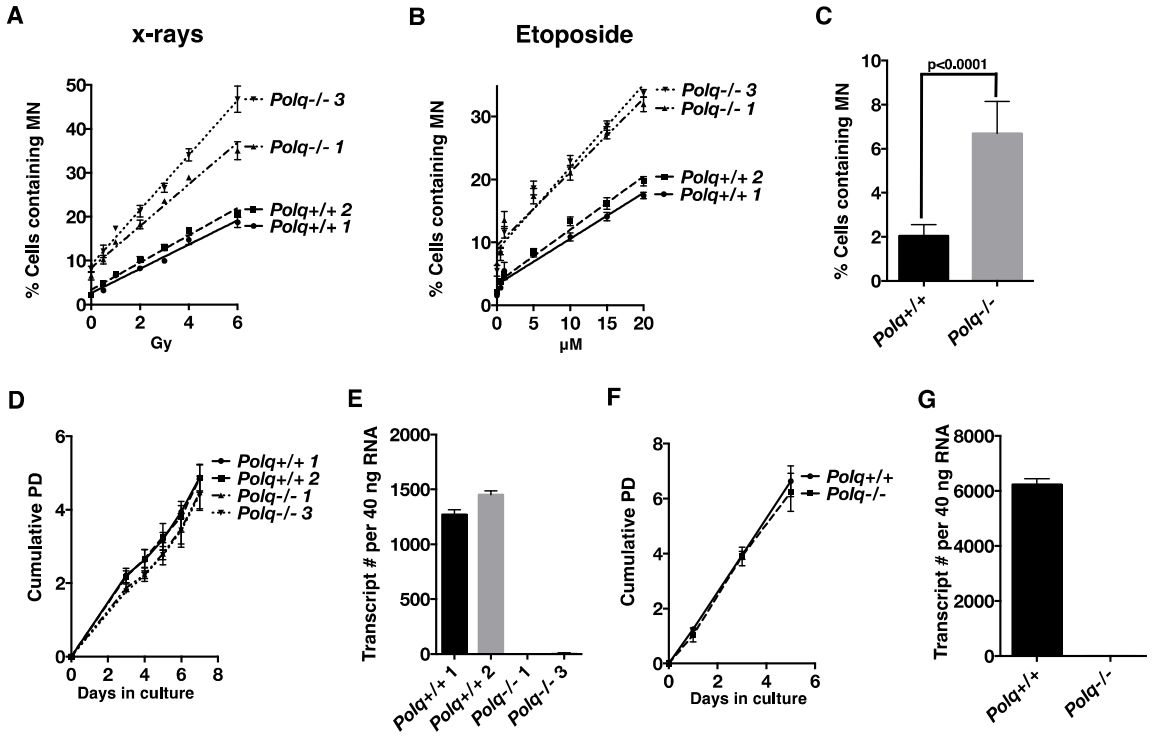


Figure 6. Loss of *Polq* contributes to chromosomal instability both spontaneously and in the presence of DNA damage. *Polq*$^{+/+}$ or *Polq*$^{-/-}$ bone marrow stromal cells (BMSCs)plated in chamber slides were exposed to (**A**) X-rays or (**B**) etoposide. 48 h after damage cells were fixed and stained with DAPI to enumerate cells with micronuclei. Counts represent the average percentage of cells with micronuclei scored in three independent experiments. [Slopes for X-ray and etoposide-induced MN: *Polq*$^{+/+}$ clone 1 (2.8, 0.73); *Polq*$^{+/+}$ clone 2 (3.1, 0.85); *Polq*$^{-/-}$ clone 1 (4.8, 1.2); *Polq*$^{-/-}$ clone 3 (6.2, 1.3)] The frequency of spontaneous micronuclei for each of the clones in **Figure 6A and B** were combined to generate (**C**) total spontaneous micronuclei observed for all genotypes. The p-value was determined by the Wilcoxon Mann Whitney rank sum test. *Polq*$^{+/+}$ or *Polq*$^{-/-}$ (**D**) BMSCs and (**F**) mouse embryonic fibroblasts were plated in growth medium in triplicate. Cells were counted at the indicated days and cumulative population doublings were recorded. The experiment was repeated three times. (**E** and **G**) Absolute quantification

of *Polq* transcript numbers in three independent experiments. Enumeration of micronuclei, cell culture, RNA isolation, cumulative population doubling analysis, and data analysis was performed by MJY. qRT-PCR was performed by the Science Park Molecular Biology Core.

The DNA polymerase activity of POLQ is required to confer resistance to DNA damaging agents

Next I sought to investigate which catalytic activities of POLQ are necessary to confer resistance to DNA damaging agents. Lentiviral-delivered expression vectors were constructed with the assistance of Dr. Junya Tomida to express wild-type or mutant versions of POLQ in immortalized *Polq*$^{-/-}$ MEFs, in order to test for functional complementation (**Figure 7A**). A tandem (D2330A, Y2331A) mutation was introduced into the DNA polymerase domain (POL) mutation of the corresponding residues in other DNA polymerases completely inactivates polymerase activity (Patel and Loeb, 2000). In a separate construct, a mutation was introduced into the conserved ATP-binding site of the Walker A motif (K121M) in the helicase-like domain (HLD). An equivalent mutation eliminates DNA helicase activity in related enzymes, including HELQ (Marini and Wood, 2002). A third construct (DM) was made harboring mutations in both domains. These vectors expressed full-length recombinant POLQ as tested in transfected 293T cells (**Figure 7B and C**).

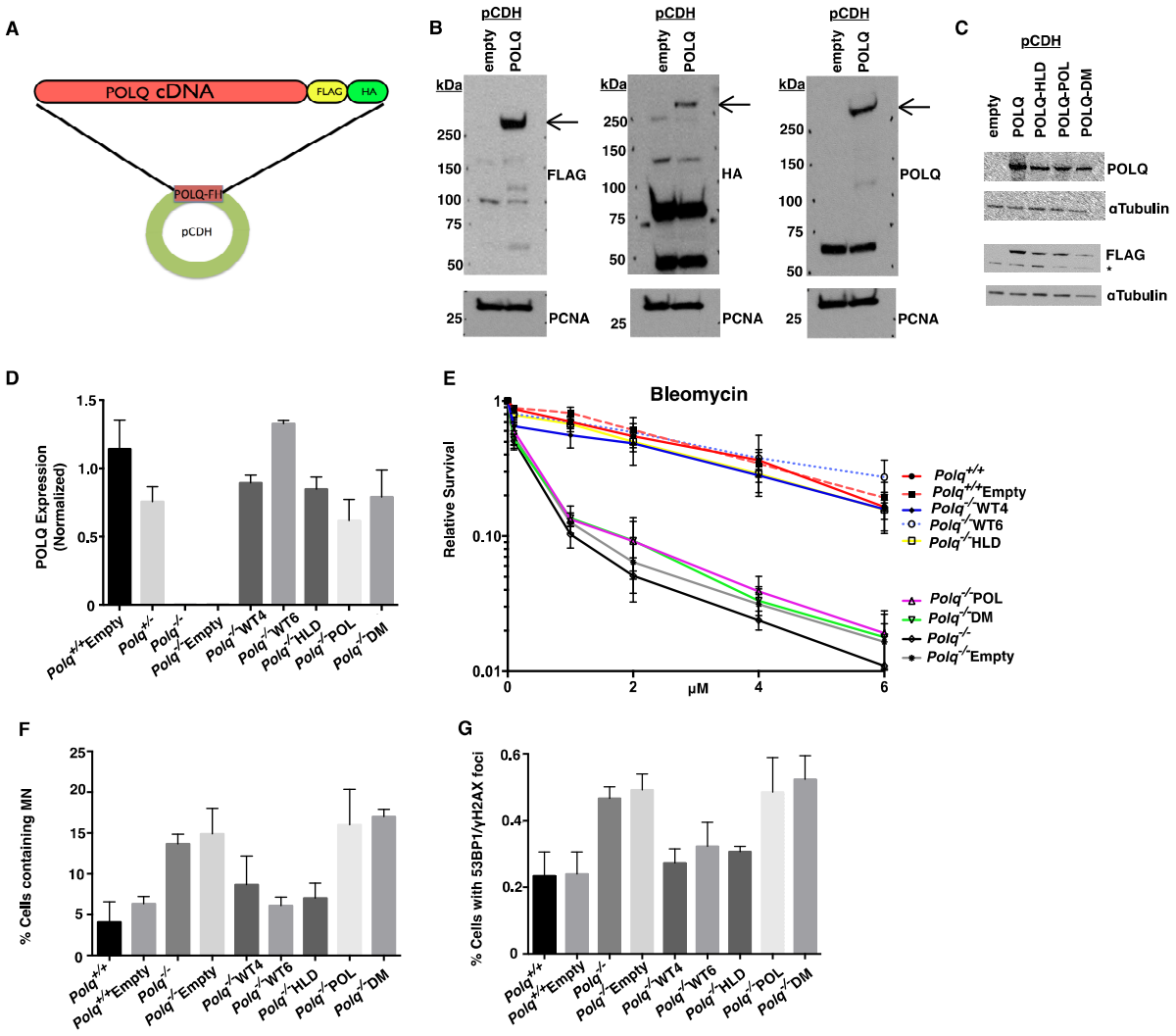


Figure 7. Complementation of the polymerase activity of POLQ rescues DNA damage hypersensitivity of cells lacking *Polq*. (**A**) POLQ cDNA was cloned into the pCDH-FH vector containing a FLAG-HA epitope tag on the C-terminus. (**B**) 293T cells were transiently transfected with pCDH containing either empty-vector control or POLQ cDNA. Crude extracts were immunoblotted with the indicated antibodies to confirm expression of full-length recombinant POLQ and (**C**) mutant constructs that were transiently transfected in 293T cells. (**D**) Stable *Polq*$^{-/-}$ MEF lines complemented with POLQ expression vectors (or empty vector control) were assayed for *Polq* expression by qPCR. WT4 and WT6 are

independent clones complemented by wild-type POLQ; POL, mutation in the DNA polymerase domain; HLD, mutation in the DNA helicase domain, DM, mutation in both domains. (**E**) The complemented MEF lines were treated with bleomycin for 24 hr and cellular ATP levels were measured 72 h later. (**F**) Spontaneous micronuclei and (**G**) DNA double-strand breaks (>2 colocalized γH2AX and 53BP1 foci per cell) were quantified for three independent experiments. The brightness of the entire microscope field was increased to better display the fluorescence for print, using Adobe Photoshop CS6. Dr. Junya Tomida generously provided the pCDH vector backbone containing the C-terminal FLAG-HA epitope tags and technical advice. All cloning of POLQ cDNA, transfections, lentiviral transductions, immunoblotting, cell sensitivity and DNA damage marker assays, and data analysis was performed by MJY. qRT-PCR was performed by the Science Park Molecular Biology Core.

The mutant cDNAs were tested for their ability to genetically complement the bleomycin sensitivity of *Polq*-null MEFs. Stable clones with each of the constructs were generated and analyzed for expression of POLQ (**Figure 7D**). Two independent clones of knockout MEFs expressing wild-type recombinant POLQ (WT4 and WT6) were able to rescue bleomycin hypersensitivity (**Figure 7E**). Neither the polymerase domain mutant (POL) nor the polymerase-helicase double mutant (DM) restored bleomycin sensitivity (**Figure 7E**, **Figure 8**). Expression of a construct with a mutation in only the helicase-like domain (HLD) was, however, still able to restore resistance to bleomycin. These data indicate that POLQ polymerase activity is essential for conferring resistance to DNA damage, whereas the ATPase activity of the helicase-like domain is not necessary. Similarly, reintroduction of POLQ polymerase activity into *Polq*-deficient MEFs was able to rescue chromosomal instability based on the analysis of micronuclei and DNA DSBs, as measured by 53BP1 and γH2AX colocalization (**Figure 7F and G**, **Figure 9**).

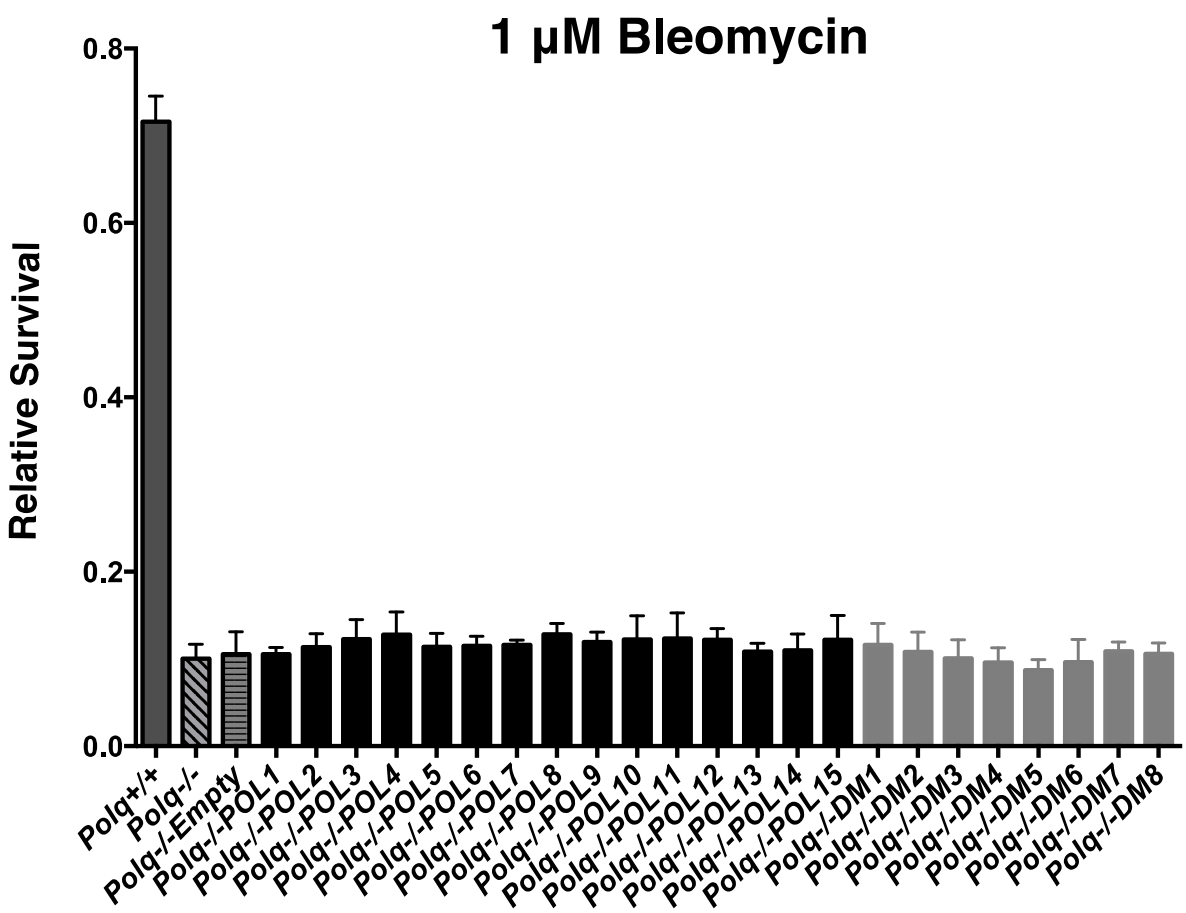


Figure 8. Complementation of polymerase-dead POLQ does not rescue DNA damage sensitivity. *Polq*$^{+/+}$, *Polq*$^{-/-}$, *Polq*$^{-/-}$Empty, and multiple clones of *Polq*$^{-/-}$POL and *Polq*$^{-/-}$DM MEF lines were treated with 1 μM bleomycin for 24 h and cellular ATP levels were measured 72 h later. All sensitivity assays were performed by MJY.

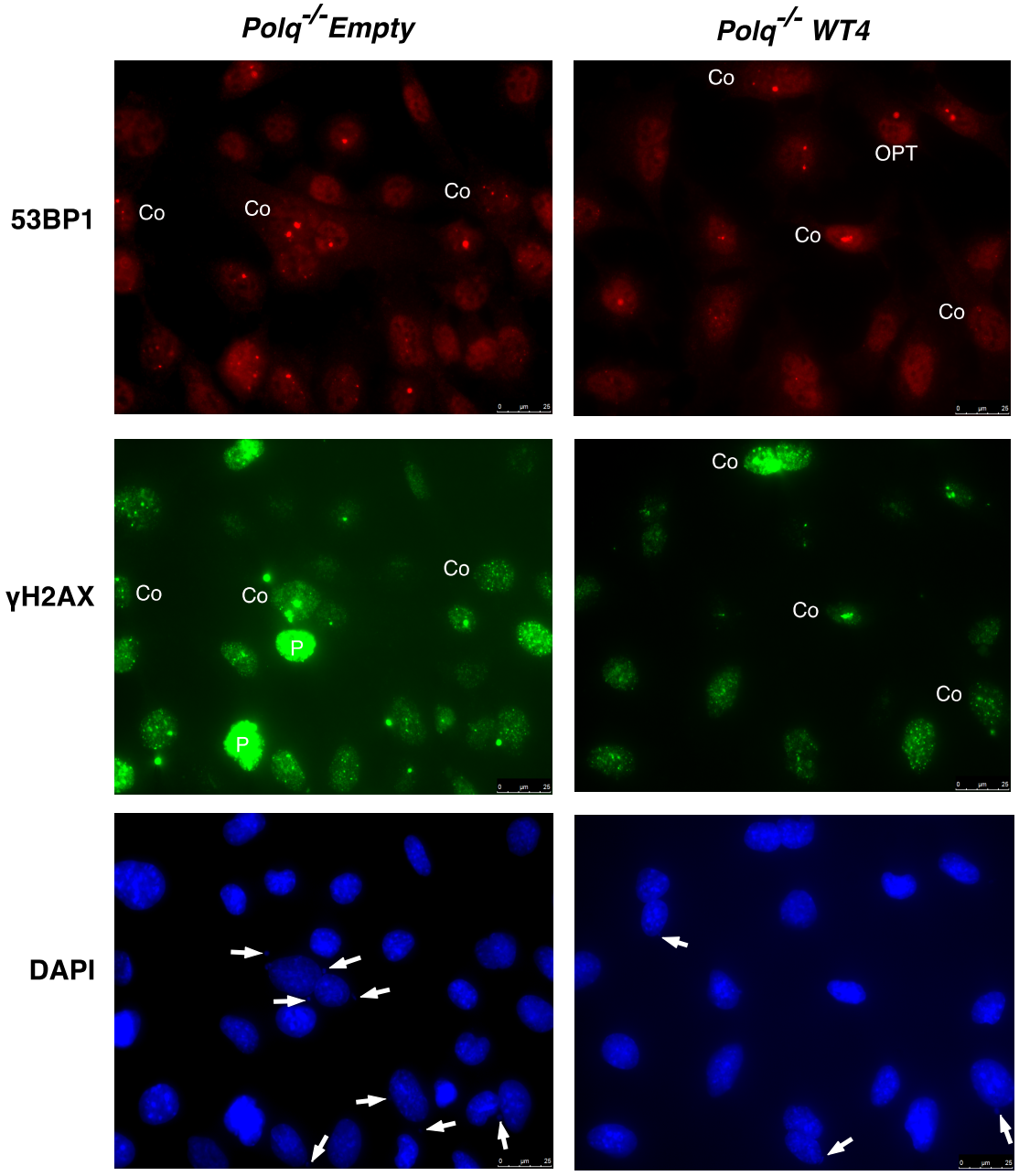


Figure 9. Analysis of DNA double-strand breaks and micronuclei in complemented *Polq* MEFs. Representative immunofluorescence images of *Polq*$^{-/-}$Empty and *Polq*$^{-/-}$WT4 MEF lines stained with DAPI and antibodies against 53BP1 and γH2AX. Scale bar represents 25 μm. Arrows note micronuclei. Pan staining of γH2AX (P), OPT domain

staining by 53BP1 (OPT), and examples of colocalization (Co) are noted. All cell culture and staining, imaging, and analysis performed by MJY.

Mice with an S1932P mutation in *Polq* (the "*chaos1*" allele) have an increased spontaneous frequency of micronuclei (Shima et al., 2003). We generated a human *POLQ* cDNA mimicking the *chaos1* mutation (S1977P), but attempted expression of POLQ with this mutation in 293T cells did not yield detectable protein despite ample transcript levels (**Figure 10**). This suggests that the *chaos1*-encoded mutant protein is unstable, consistent with the finding that *chaos1* mice have a phenotype essentially indistinguishable from *Polq* knockout mice (Shima et al., 2003).

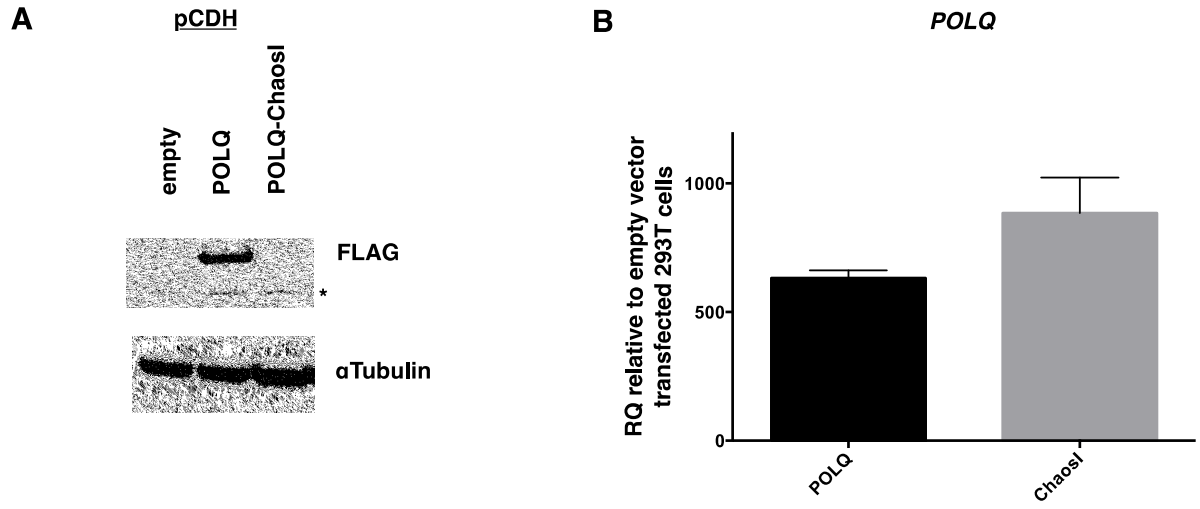


Figure 10. Full-length *chaosI* mutant protein is poorly expressed. 293T cells were transfected with pCDH constructs that contained either the human POLQ cDNA, *ChaosI* mutant (S1977P, corresponding to the S1932P mutated residue in mice), or empty control. (**A**) 48 h post transfection lysates were made and immunoblotted with antibodies against FLAG and alpha-Tubulin. (**B**) Total RNA was isolated from cells and qPCR analysis of

POLQ transcript levels were performed using the ΔΔCt method. An asterisk (*) denotes a non-specific band. All cell culture, transfections, immunoblotting, RNA isolation, and data analysis were performed by MJY. qRT-PCR was performed by the Science Park Molecular Biology Core.

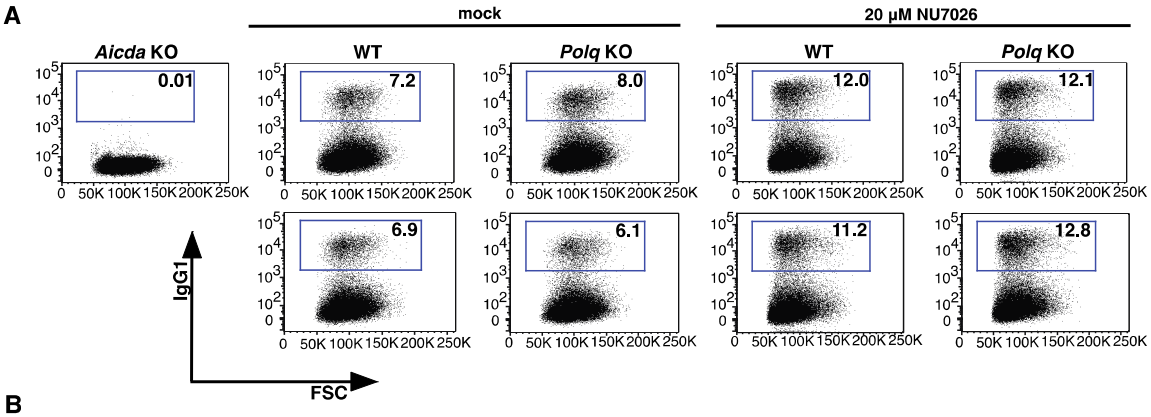
POLQ operates in a pathway of altEJ during mouse Ig class-switching

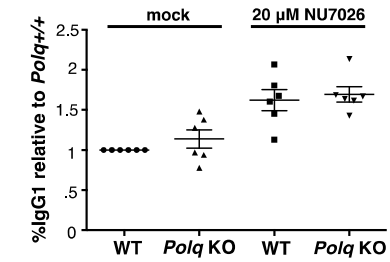
Immunoglobulin class-switch recombination (CSR) uses DNA end joining to exchange one constant region of an antibody gene for another constant region (**Figure 4**). CSR can occur by both Ku-dependent classical non-homologous end joining and Ku-independent altEJ (Boboila et al., 2012a). The overall frequencies of CSR are similar in *Polq*-defective mice (Martomo et al., 2008) and cultured B cells (Li et al., 2011). To determine whether POLQ is involved in a mechanistically distinct subset of CSR joins, In 2012 our lab began a collaboration with Dr. Kevin McBride, a molecular immunologist on the department faculty, who had expertise in CSR. I, with help from Dr. Kevin McBride and his laboratory, isolated and analyzed DNA sequences at such joins. Naïve B cells were isolated from the spleens of wild-type and *Polq*-null mice and stimulated for IgM to IgG class switching, and then the fraction of IgG1-positive B cells was measured by flow cytometry. Parallel B-cell cultures were incubated with NU7026, a DNA-PKcs inhibitor that suppresses cNHEJ (Callen et al., 2009). It has been shown that B cells incubated with NU7026 have an increased proportion of CSR junctions with >1 bp insertion at the junction (Callen et al., 2009). This suggests that when altEJ is used during CSR, it more frequently results in the insertion of nucleotides.

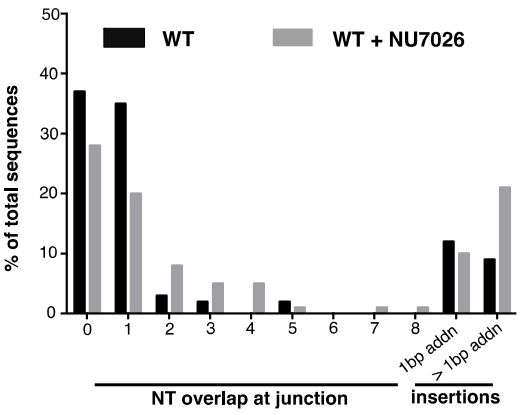
I found that B cells from *Polq*-proficient and deficient mice had similar overall frequencies of CSR (**Figure 11A**), and that inhibition of DNA-PKcs increased the frequency of CSR in both genotypes by 1.5 to 2-fold (**Figure 11B**). The Sμ-Sγ1 junction was then

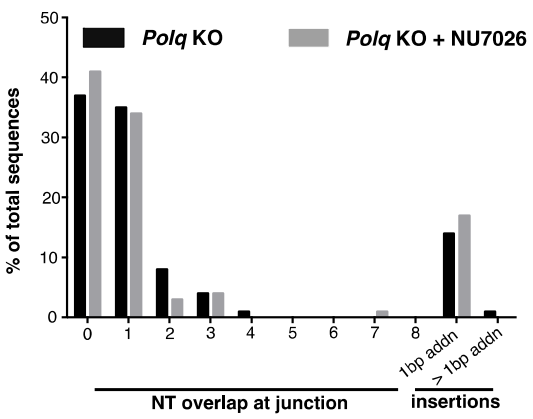
sequenced from 100 clones of each group of IgG1-positive B cells. In wild-type B cells, insertions of >1 bp at Sμ-Sγ1 junctions, that are thought to be altEJ-dependent, comprised about 9% of total events, and was increased to ~21% in cells incubated with NU7026 (**Figure 11C**, **Table 4**).

Strikingly, in cells lacking *Polq*, this class of longer insertions at CSR junctions was absent, even in the presence of NU7026 (**Figure 11D**), only one insertion of 2 bp was observed, therefore insertions of >1 bp require POLQ. *Polq*-dependent joining events included insertions of between 2 and 35 bp. For longer insertions (greater than ~10 bp) homologous sequences were unambiguously detected up to 2–5 kbp away from the junction site (**Table 4**), as has been reported for long insertions at Sμ-Sγ1 junctions in ATM-defective B cells (Callen et al., 2009). This suggests that most, if not all, such insertions are formed in a templated manner during altEJ by POLQ.


A
mock
20 μM NU7026
Aicda KO
WT
Polq KO
WT
Polq KO
0.01
7.2
8.0
12.0
12.1
6.9
6.1
11.2
12.8
IgG1
FSC


B
mock
20 μM NU7026
%IgG1 relative to Polq+/+
WT
Polq KO
WT
Polq KO


C
WT
WT + NU7026
% of total sequences
NT overlap at junction
insertions
1bp addn
>1bp addn


D
Polq KO
Polq KO + NU7026
% of total sequences
NT overlap at junction
insertions
1bp addn
>1bp addn

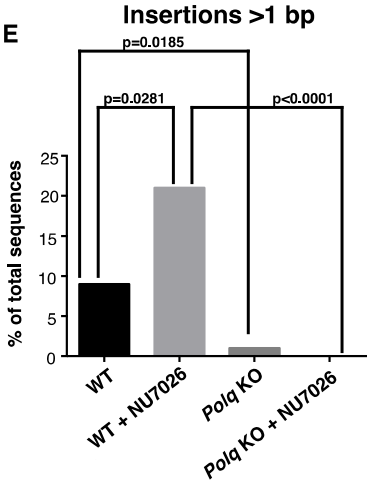

E
Insertions >1 bp
p=0.0185
p=0.0281
p<0.0001
% of total sequences
WT
WT + NU7026
Polq KO
Polq KO + NU7026

Figure 11. Insertions >1 nt at CSR junctions are *Polq*-dependent. Isolated wild-type (WT) and *Polq*$^{-/-}$ (KO) naïve splenic B cells were stimulated for CSR and either mock-treated or treated with NU7026. (**A**) Cells were assayed for IgG1 levels (y-axis) by flow cytometry. The x-axis displays forward scatter (FSC) for cell viability. *Aicda*$^{-/-}$ splenic B cells were used as a negative control since they do not undergo CSR. Numbers in boxes show the percentage of the population that is IgG1 positive; in (**B**) these data are plotted relative to wild-type. Genomic DNA isolated from B cells of wild-type (**C**) and *Polq* KO (**D**) mice was amplified by PCR and 100 Sμ-Sγ1 junctions from each group were sequenced and analyzed for overlaps and insertions at breakpoints. (**E**) Insertions >1 nt are plotted; p-values were determined by the two-tailed Fisher's exact test. Cell viabilities were comparable between genotypes. B cell isolation and culture was performed by MJY with assistance from Dr. Kevin McBride and Sean Hensley. FACS analysis was performed by Dr. Kevin McBride, and DNA isolation, PCR amplification, and sequencing analysis of switch junctions was performed by MJY. DNA sequencing was performed by the Science Park Molecular Biology Core.

	Sμ (J00440)	Insertion	Sγ (D78344)		Insert Homology
	Polq$^{+/+}$				
4648-4677	GGGACCCAGGCTTTGAAGGCAATCCTGGGA	GCACTTC	GTACTTATAGAGGAACAGGGGCAGCGTAGA	2888-2917	
5161:5190	ACTGTAATGAACTGGAATGAGCTGGGCCGC	CCT	GGAAATCAGGACAGGTACAAGCGTGTGGAT	7925:7954	
5291:5320	TTCTGAGCTGAGATGAGCTGGAGTGAGCTC	TTACC	GGTAGCTATAGGGAATACAGAACAGGTGGG	2653:2682	
4974:5003	GAAGGCGAGACTCATAAAGCTTGCTGAGCA	CCAGGACA	CCCTGTAGGGTAGCTGTAGGGAAATCAGGA	7906:7935	Sγ2278:2285
5206:5235	TGGCTTAACCGAGATGAGCCAAACTGGAAT	CCC	GAATAGCTACAGGGGAGCCAGGAGAGGTAG	2753:2782	
4891:4920	AATTGAGAAAGAATAGAGACCTGCAGTTGA	TC	CTGATGGCAAATGGAAGGGCAGAGACCCAG	2957:2986	
5110:5139	AGGTGATTACTCTGAGGTAAGCAAAGCTGG	CA	CCAGGTCTGCAGCTACATACGGGTAAGCAG	2694:2723	
5297:5326	GCTGAGATGAGCTGGGGTGAGCTCAGCTAT	TT	TATAGGGAGCCAGGACAGGTGGAAGTGTGG	7602:7631	
4667:4696	CAATCGGGGGATTCTGGAAGAAAAGATGTT	CC	GGATCAAGGCAGAACAGGTCCAGGGGTGCC	2840:2869	
	Polq$^{+/+}$ + NU7026				
5432:5461	TAGGGTGAGCTGAGCTGGGTGAGCTGAGCT	CCCAGT	AGTGTGGGGATCCAGGTAAGGCTGGACTGG	8264:8293	
4620-4649	TGGAAGCTAATTTAGAATCAAGTAAGGAGG	CAAGACAGGTGG--GTGTGGGGATCAAGG	AGGAGAAATGGAAGAATGCAGATTCCAAAC	8030:8059	Sγ2820:2848
4907:4936	AGACCTGCAGTTGAGGCCAGCAGGTCGGCT	ACTA	TGGAACTGTGGGGACCCTGTAGGGTAGCTG	7892:7921	
4787:4816	GTTGTTAAAGAATGGTATCAAAGGACAGTG	CTAGAACGTACT	TCAAGGCAGAACAGGTCCAGGGGTGCCAGG	2843:2872	Sμ3545:3556
5231:5260	GGAATGAACTTCATTAATCTAGGTTGAATA	GGTGGAAGTGTGGGG	AGGGCAGCCAGGACAGGTGGAAATGTGGTG	7750:7779	Sγ4882:4896
4879:4908	ACAGCTGTACAGAATTGAGAAAGAATAGAG	TGGGG--AGCTCAGCTATGCTAC-CGTGTTG-GG	CAGGAAGAATAGCTACAGGGGAGCCAGGAG	2747:2776	Sμ5309:5342
5291:5320	TTCTGAGCTGAGATGAGCTGGAGTGAGCTC	TTACC	GGTAGCTATAGGGAATACAGAACAGGTGGG	2653:2682	
4974:5003	GAAGGCGAGACTCATAAAGCTTGCTGAGCA	CTACGACA	CCCTGTAGGGTAGCTGTAGGGAAATCAGGA	7906:7935	
5237:5266	AACTTCATTAATCTAGGTTGAATAGAGCTA	GAC--GTAGCATTGTGTGACTC	CAGGCACAGTAGCTATAGGGGAGCCAAGAC	2796:2825	Sγ2575:2596
4989:5018	AAAGCTTGCTGAGCAAAATTAAGGGAACAA	TTGAG	TGAGGCAGGTAAGAGTGTGGGAACCCAGTC	8429:8458	
5245:5274	TAATCTAGGTTGAATAGAGCTAAACTCTAC	AGTGCAGA	TGAGGCAGGTAAGAGTGTGGGAACCCAGTC	8429:8458	Sγ8821:8828
4685:4714	AGAAAAGATGTTTTTAGTTTTTATAGAAAA	GTAAT	TATAGGGGAGCCAAGAAAGGTGGAAGTGTG	2809:2838	
4796:4825	GAATGGTATCAAAGGACAGTGCTTAGATCC	GAGGTGAGTGTGAGAGGACAAA	AAGTTTAGTAGTTATAGAGGAACAGGGGCA	2881:2910	Sμ4827:4845
5088:5117	GCTTCTAACATGCGCTAAACTGAGGTGATT	CCAAGGACCCAGGCAGAGCAGCTCCAGTAGGCCA	CAATCACAAGGGAACTGATGGCAAATGGAA	2943:2972	Sγ4014:4035
4663:4692	AAGGCAATCTTGGGATTCTGGAAGAAAAGA	CCAT	GAAGTGTAGTGACCCTTGAAGAATAGCTAC	2733:2762	
5002:5031	CAAAATTAAGGGAACAAGGTTGAGAGCCCT	TC	GGAATTGTGGTGACCCAGACAAAACAGCTA	7817:7846	
4684:4713	AAGAAAAGATGTTTTTAGTTTTTATAGAAA	CC	ACAGGGAAGCTATATTAAAACCAGGACACG	8132:8161	
5133:5162	AAGCTGGGCTTGAGCCAAAATGAAGTAGAC	GC	TGGAAGGGCAGAGACCCAGACTAAATGGCT	2968:2997	
5210:5239	TTAACCGAGATGAGCCAAACTGGAATGAAC	TA	AAAACCAGGACAGGAGGAAGAATGGGGATC	8148:8177	
4816:4845	GCTTAGATCCGAGGTGAGTGTGAGAGGACA	AA	AAGTTTAGTAGTTATAGAGGAACAGGGGCA	2881:2910	
4897:4926	GAAAGAATAGAGACCTGCAGTTGAGGCCAG	TT	AGGGTGTGGATCCAGGCAGGGTAGCTATAG	2634:2663	
	Polq$^{-/-}$				
5360:5389	AGCTACTCTGGAGTAGCTGAGATGGGGTGA	TT	CAAATGGAAGGGCAGAGACCCAGACTAAAT	2964:2993	

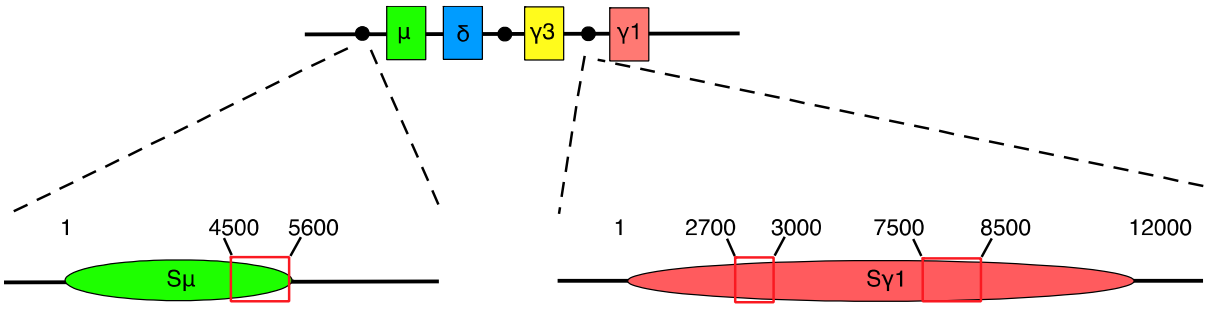


Table 4. Sequence composition of >1 nucleotide CSR insertions in *Polq* mice. New mutations (different from the reference sequence) are shown in red. For some longer insertions (blue), homologies were identified in the switch region, at the positions indicated in the right column. Microhomologies at the junction site are underlined. Sequences (graphed in **Figure 11**) are from wild-type (*Polq*$^{+/+}$) and *Polq*-null (*Polq*$^{-/-}$) splenic B cells that were treated with NU7026 or mock-treated. Schematic below table shows CSR mu (Sμ) and gamma 1 (Sγ1) switch regions with areas of DNA breakage noted by red boxes. Numbers correlate to switch region length. Sequence analysis performed by MJY.

Loss of Polq impairs an altEJ pathway but not cNHEJ in cultured cells

The most important factor in determining which double-strand break repair pathway is used is whether or not the 5′ termini of broken ends are resected (Symington and Gautier, 2011). Ends with little or no single stranded overhang are typically rejoined by Ku-dependent cNHEJ. In contrast, CtIP and MRN-dependent resection of 5′ termini generates ends with extended single stranded 3′ overhangs; resection is thought to block cNHEJ (Frank-Vaillant and Marcand, 2002) and enable repair by altEJ (Bennardo et al., 2008; Lee and Lee, 2007).

In 2013, Dr. Dale Ramsden (UNC Chapel Hill) contacted Dr. Wood to enquire about POLQ and DNA end joining. We initiated collaboration where I sent my complemented *Polq* MEF lines to be used to determine the impact of POLQ status on cNHEJ and altEJ. To analyze differing requirements for end joining, with or without end resection. David Wyatt, a graduate student in the Ramsden laboratory, prepared two linear DNA substrates with 3′ single stranded overhangs; one with a short overhang (6 nt), and one a long overhang (45 nt, a “pre-resected end”) (**Figure 12A**). Both can be aligned with the same 4 nt of terminal complementary sequence. These substrates were then introduced into wild-type mouse fibroblasts or fibroblasts harboring deficiencies in *Ku70* or *Polq*. Repaired products were recovered from cells and quantified. Repair of the short overhang substrate was, as anticipated, over 10-fold less efficient in cells without Ku70 (**Figure 12B**) when compared to Ku70-complemented controls. The absence of *Polq* had no consequence for the repair of this substrate.

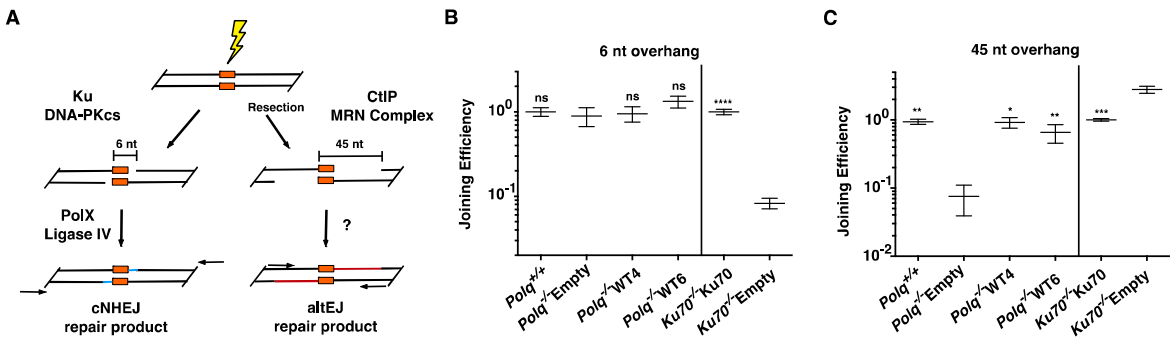


Figure 12. End joining with extrachromosomal substrates. (**A**) Substrates were designed to resemble DNA double-strand breaks that are repaired through Ku-dependent NHEJ (6 nt tail with 4 nt of terminal complementary sequence) or alternate end joining of resected DNA substrates (45 nt tail with 4 nt terminal complementary sequence), introduced into cells, and joining of head-to-tail products assessed by qPCR. (**B**) qPCR for the classical NHEJ assay uses primers to detect all events having sequences in the duplex immediately flanking the break. Joining efficiencies are expressed as fractions of the mean joining determined for matched wild-type controls (*Polq*$^{+/+}$ or Ku70 complemented lines, as appropriate). Three independent triplicate measurements were made for the *Polq* cell lines and two independent triplicates for the *Ku* cell lines. Error bars represent the standard error of the mean. Joining efficiency was not significantly different, whether cells were deficient in *Polq*(*Polq*$^{-/-}$Empty) or not (*Polq*$^{+/+}$, *Polq*$^{-/-}$WT4, *Polq*$^{-/-}$WT6), but was different when cells expressed Ku (*Ku70*$^{-/-}$Ku70) when compared to *Ku70*$^{-/-}$Empty cells (t-test, $p<0.05$) (**C**) qPCR for the altEJ assay used primers to detect that subset of products including at least 10 nt of each 3′ overhang. Mean relative joining efficiencies, standard error of the mean, and statistical analysis performed as for panel B. Joining efficiency was significantly different in cells expressing *Polq* (*Polq*$^{+/+}$, *Polq*$^{-/-}$WT4, or *Polq*$^{-/-}$WT6) when compared to *Polq*$^{-/-}$Empty cells ($p<0.05$), and in cells expressing Ku (*Ku70*$^{-/-}$Ku70) when compared to *Ku70*$^{-/-}$Empty cells (t-test, $p<0.05$). The background observed in a mock transfected sample was

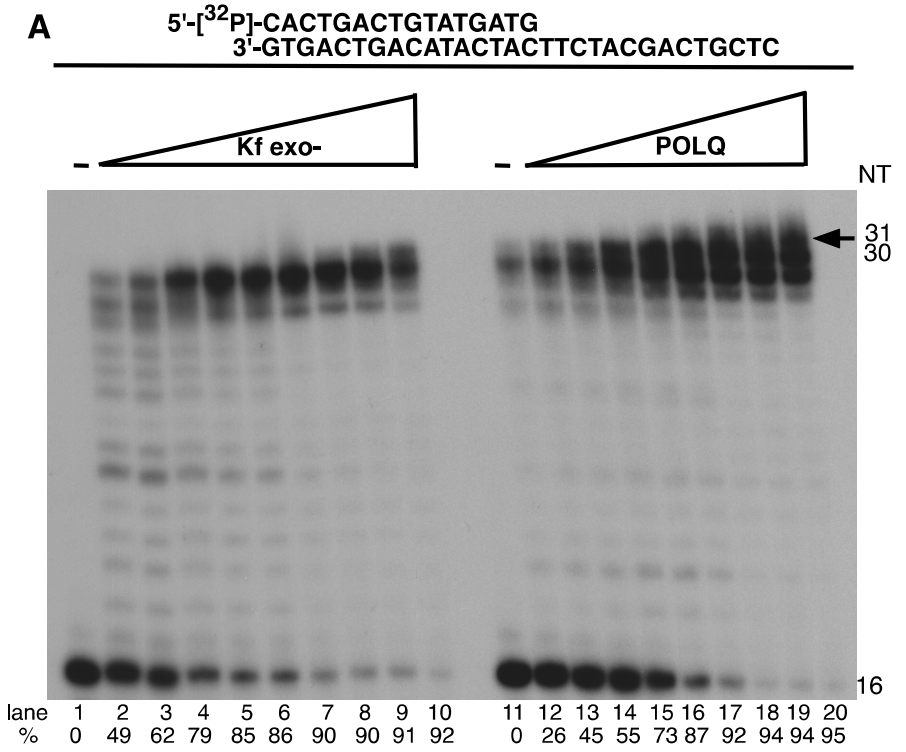
determined to be 0.038, +/– 0.02 of wild-type controls. p values are represented as: * p<.05, ** p<.01, *** p<.001, ****p<.0001, and not significant (ns). Substrates were designed by David Wyatt and Dr. Dale Ramsden, assay performed by David Wyatt, and data analyzed by David Wyatt and Dr. Dale Ramsden. *Polq*-null and complemented cell lines were generated and provided by MJY

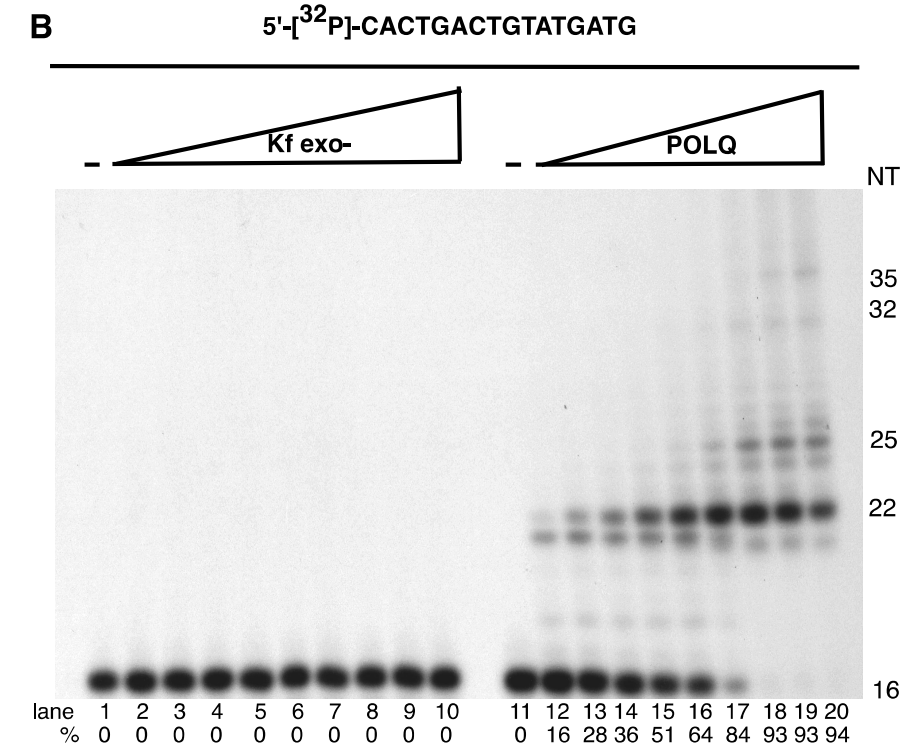
End joining with the 45 nt overhang substrate was assessed using qPCR primers located to ensure that at least 10 nt of overhang was included in joined products (**Figure 12A**). Recovery of these products was no longer dependent on Ku; instead, it was increased 2.8-fold in *Ku70*-deficient cells (**Figure 12C**). This is consistent with previous studies demonstrating that Ku suppresses repair by altEJ. Strikingly, joining of the long overhang substrate in *Polq*$^{-/-}$ cells was reduced 10-fold, approaching background levels of signal observed using this assay. Complementation of the knockout cells with POLQ returned altEJ joining to wild-type levels (**Figure 12C**). These data demonstrate that POLQ participates in some form of altEJ, but that cells lacking POLQ maintain proficiency for cNHEJ.

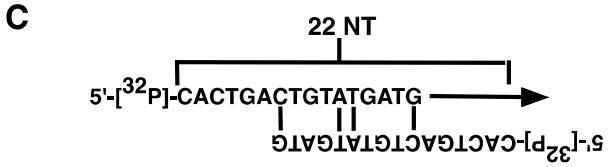
POLQ extends 3′ DNA ends in a template-dependent manner

Our results demonstrate that POLQ is necessary to form the insertions found in CSR junctions in a process of altEJ. We next sought to determine the mechanism. Dr. Kei-ichi Takata, an expert in DNA polymerase assays in the Wood laboratory, generously provided technical advice and performed primer extension analysis to further analyze the polymerase properties of POLQ *in vitro*. Like other DNA polymerases, an active polymerase fragment of POLQ (Hogg et al., 2011) can catalyze template-dependent DNA synthesis from an annealed primer (**Figure 13A**). As is common for family-A DNA polymerases, only a single nucleotide is added to the end of duplex DNA (Seki et al., 2004). Unusually, however, POLQ

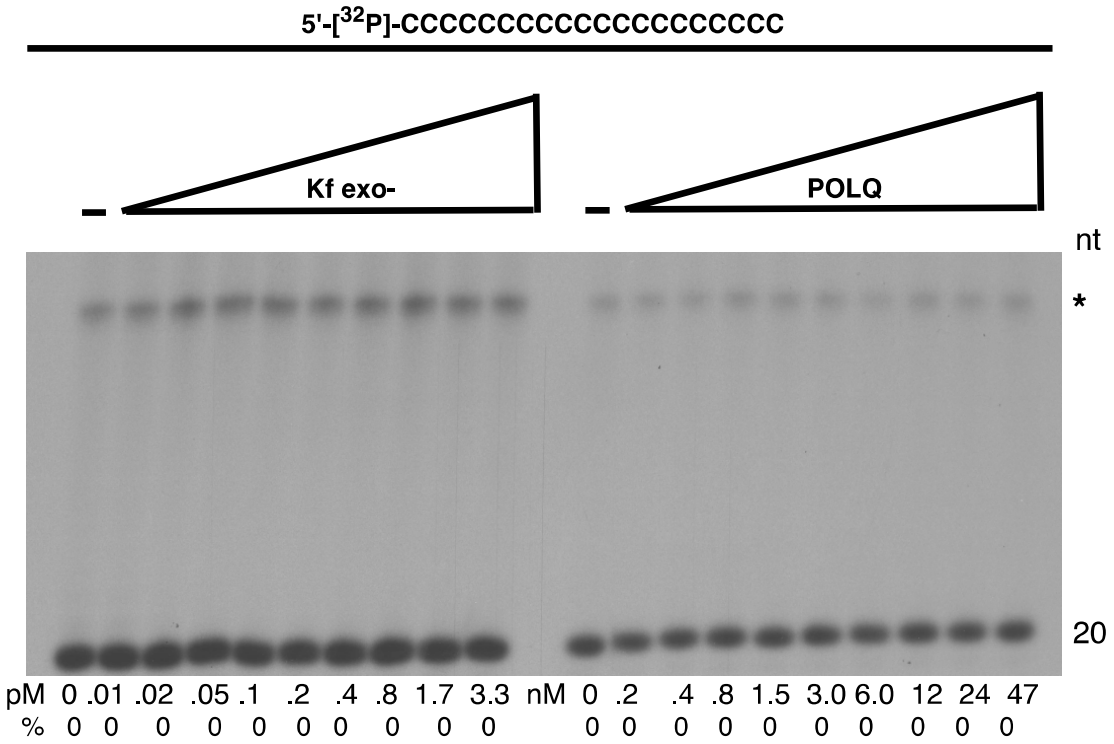
can catalyze extension of single-stranded oligonucleotides (Hogg et al., 2012). It was unclear whether this reflects a robust terminal deoxynucleotide transferase activity of POLQ on single-stranded DNA, or some form of template-dependent synthesis. For example, POLQ can extend a single-stranded 16-mer oligonucleotide provided without a complementary template (products up to 35 nt long), while *E. coli* pol I Klenow fragment has no activity on this substrate (**Figure 13B**). The major 22 nt extension product produced by POLQ on the 16-mer used in **Figure 13B** may be accounted for by inter- or intra-oligonucleotide pairing (**Figure 13C**). Neither POLQ nor Klenow fragment could extend an oligonucleotide that was incapable of annealing to itself (**Figure 13D and E**) (Hogg et al., 2012).


A
5'-[32P]-CACTGACTGTATGATG
3'-GTGACTGACATACTACTTCTACGACTGCTC
Kf exo-
POLQ
NT
31
30
16
lane 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
% 0 49 62 79 85 86 90 90 91 92 0 26 45 55 73 87 92 94 94 95


B
5'-[32P]-CACTGACTGTATGATG
Kf exo-
POLQ
NT
35
32
25
22
16
lane 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
% 0 0 0 0 0 0 0 0 0 0 0 16 28 36 51 64 84 93 93 94


C
22 NT
5'-[32P]-CACTGACTGTATGATG
5'-[32P]-CACTGACTGTATGATG

D


5'-[32P]-CCCCCCCCCCCCCCCCCCCC
Kf exo-
POLQ
nt
*
20
pM 0 .01 .02 .05 .1 .2 .4 .8 1.7 3.3 nM 0 .2 .4 .8 1.5 3.0 6.0 12 24 47
% 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

E

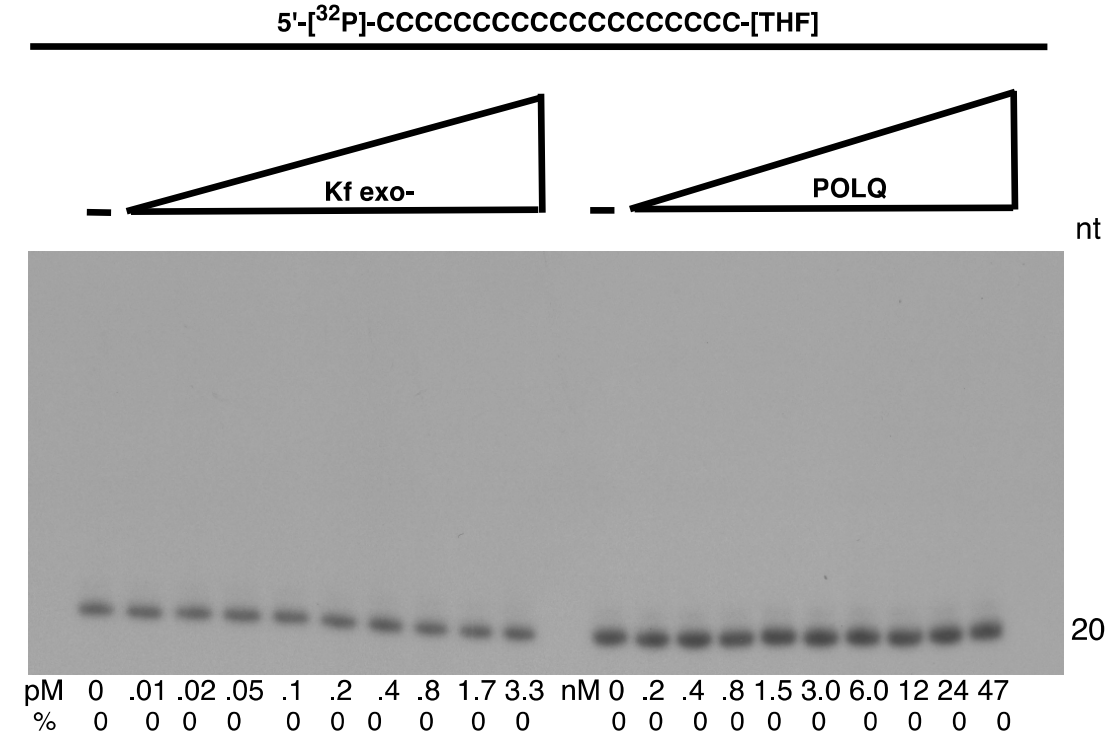

5'-[32P]-CCCCCCCCCCCCCCCCCCCC-[THF]
Kf exo-
POLQ
nt
20
pM 0 .01 .02 .05 .1 .2 .4 .8 1.7 3.3 nM 0 .2 .4 .8 1.5 3.0 6.0 12 24 47
% 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

Figure 13. Unique template dependent DNA polymerase activity of POLQ.

Exonuclease-defective *E. coli* pol I Klenow fragment (Kf exo-) or POLQ was incubated at the indicated protein concentrations with (**A**) a 5′-^{32}P-labeled 16-mer primer and 30-mer complementary template, or a (**B**) 5′-^{32}P-labeled 16-mer primer and no template. All reaction mixtures included all four deoxynucleotide triphosphates and were incubated at 37°C for 10 min (**A**) or 20 min (**B**). The first lane contained no enzyme. The percentage (%) of the primer extended is shown below each lane. (**C**) Model of major 22 nt product formation produced by extension from the primer in **Figure 13B**, which has a limited ability to self-anneal. Similar reactions were performed with (**D**) 5′-^{32}P-labeled 20-mer dC oligonucleotide or (**E**) a 20-mer dC oligonucleotide with a synthetic abasic site on the 3' end. Primer extension assays were performed by Dr. Kei-ichi Takata and MJY.

In the summer of 2013 at a meeting in St. Petersburg, Russia, Dr. Erik Johansson (Umeå University, Umeå, Sweden) approached Dr. Wood with some very interesting biochemical data about POLQ. To identify the mechanism of 3′ single-stranded DNA extension by POLQ, Dr. Göran Bylund and Sara Lundsten of the Johansson laboratory. They used a single-stranded oligonucleotide designed to be unable to form self-complementary base pairs longer than a single nucleotide (Hogg et al., 2012), and sequenced the products of POLQ-mediated extension. Individual extension products of 1 to 30 nt were found (**Table 5**). Most of the sequenced extension products feature AAC or AAAC sequences that could arise from copying GTTT sequences in the template via inter- or intra-molecular priming and re-priming (**Figure 14**) following minimal base pairing at the 3′-primer end. My analysis of these data reveal that POLQ uniquely extends 3′ DNA tails through template-dependent DNA synthesis from a primer with minimal base pairing rather than through a true TdT-like activity. POLQ indeed has unique biochemical properties

compatible with these observations. Unlike other DNA polymerases, POLQ can efficiently extend a DNA chain with a nucleotide incorporated opposite an abasic site (Seki et al., 2004), or from a mismatched primer-terminus (Seki and Wood, 2008). Further, there is evidence that primers slip on DNA templates with an increased frequency during POLQ-mediated synthesis, as shown by the high frequency of single base pair frameshift mutations generated by purified POLQ (Arana et al., 2008).

Substrate **GTTTGTGAGTTTCC**

Seq	Length	Extension Product	Homopolymer
1	30	AAACAAACAAACAAACAAACAAACAAACAA	T
2	26	AAAACAAGCAAACAACCAAACTTTTG	T
3	25	AAAAAAAAAAACAAACTCACGCGAT	G
4	24	AACACAAACAAAACAACCCGCAAT	C
5	23	ACAAACAAACAAACAAACACACG	C
6	21	AAAACAACAAACAAAACTCCG	C
7	21	AAACAAAACAAACAAACAAAG	A
8	20	AAACAAACAAACAAGCCGGG	C
9	20	AACAAACCTGTACCGGCCCG	C
10	20	TCAAAAAACTCAAAACAAAC	A
11	19	AAAACAAAACAAAACAAAC	G
12	19	CAAACTCATCTCACGCGAA	C
13	18	AAACAAACAAAACAAACC	A
14	18	ACAAACAAACAAACACAC	A
15	17	AACAAACAAAACAAACA	C
16	17	ACAAACAAACAAACCGG	C
17	17	AAAAACAAAAAAAACTC	A
18	16	AAGCAAACAACAAAAA	G
19	16	AAAAACAAAAAAAAAA	G
20	14	AAAACAAAACAAAA	G
21	14	AAACAAACAAACAC	G
22	14	ACAAACAAACAAAC	A
23	13	AAACAAACAAACG	A
24	13	AAACAAACAAGAG	A
25	13	AAACAAACAAACG	A
26	12	ACAAAACAAACG	C
27	12	AAACAAACAAAC	A
28	12	AAACAAACAAAC	A

Substrate **GTTTGTGAGTTTCC**

Seq	Length	Extension Product	Homopolymer
29	11	AAACAACCAAC	A
30	11	ACAACCCCAAC	A
31	9	AAACAAACG	A
32	8	GGGCGAGT	G
33	8	AAACAAAC	A
34	8	AACAAGAC	A
35	6	GCGCCG	C
36	5	AAAAC	A
37	5	AAGAC	A
38	5	CACCC	A
39	5	CACAC	A
40	5	AAGAC	A
41	4	TCGT	G
42	4	GAGG	A
43	4	AAGC	A
44	3	AAA	G
45	3	CAC	A
46	2	AA	G
47	2	CC	A
48	1	A	G
49	1	T	G
50	1	T	G
51	1	T	G
52	1	C	G
53	1	T	A
54	1	C	A

Table 5. POLQ-dependent extension of 3' DNA ends relies upon homology. Primer extension assays were carried out with a defined substrate (top) incubated in the presence of an active polymerase fragment of POLQ. Products were then incubated with terminal deoxynucleotidyl transferase (TdT) in the presence of only dA, dC, dG, or dTTP to create extension products that terminate with homopolymeric runs. The reaction products were then cloned and sequenced for analysis of extension products. Extension products are listed

above for each homopolymeric run. Polymerase assay and sequencing of extension products were performed by Dr. Göran Bylund and Sara Lundsten. Data analysis was performed by MJY to make the model in **Figure 14**.

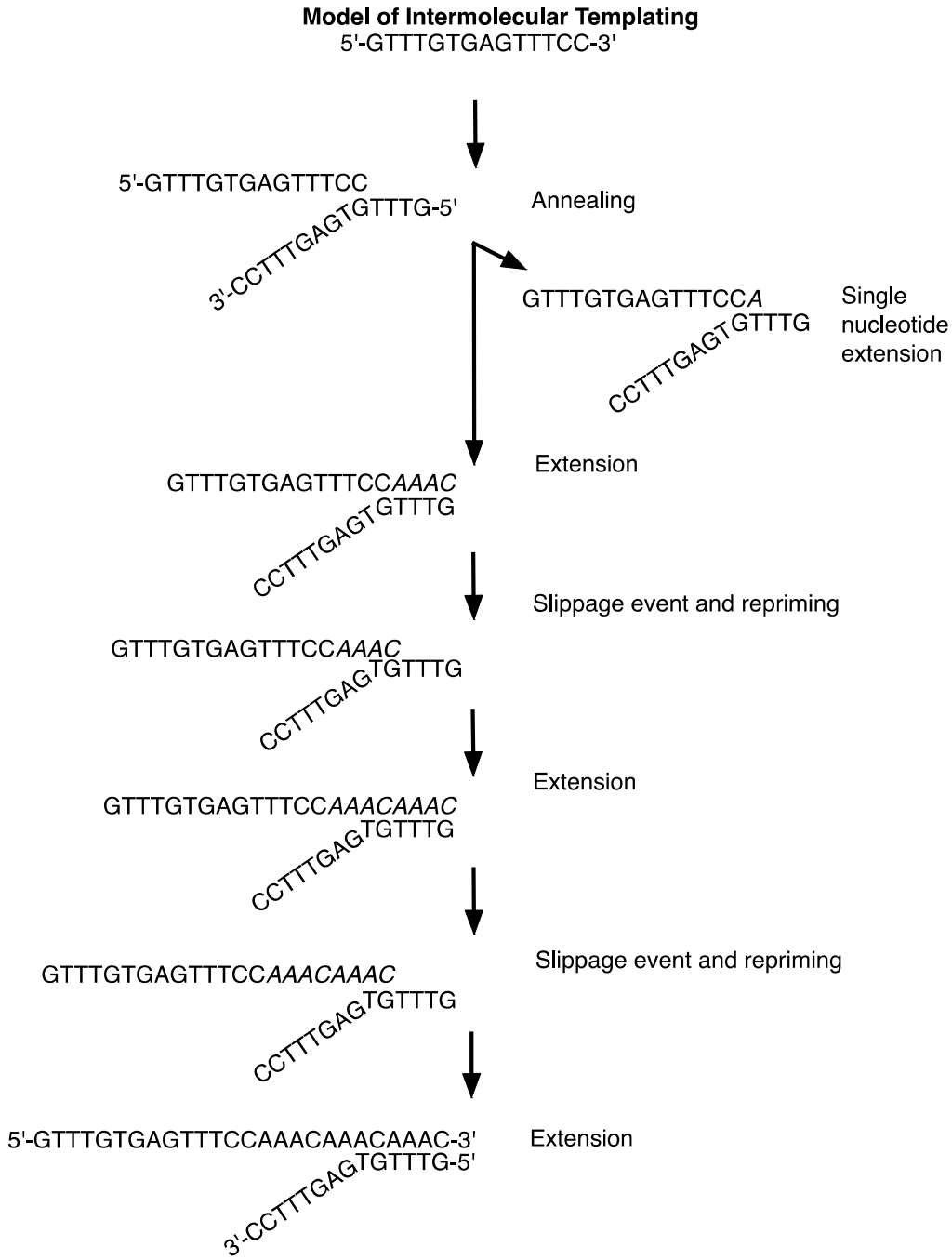


Figure 14. POLQ promiscuously extends single-stranded DNA ends through intermolecular templating. Model of intermolecular templating performed by POLQ in the process of extending a single-stranded oligonucleotide, used to produce the data in **Table 5**. This model depicts a 12 nt extension product in **Table 5**. The product can be produced by a series of annealing, extension, slippage and repriming events. DNA ends could also be elongated, in some cases, through intramolecular templating, extension, followed by slippage.

POLQ suppresses Myc/IgH translocations in B cells

Double-strand breaks initiated by AID activity in the immunoglobulin heavy chain (IgH) locus of B cells are necessary to generate immunological diversity, but breaks are sometimes generated at other chromosomal sites, providing an opportunity for dangerous chromosome translocations (Chiarle et al., 2011; Klein et al., 2011; Simsek and Jasin, 2010; Zhang and Jasin, 2011). For instance the oncogenic *Myc/IgH* translocation that causes Burkitt lymphoma is AID-dependent and requires breaks at both *Myc* and *IgH*, with breaks in the *Myc* gene being rate-limiting (Robbiani et al., 2008). An altEJ process is implicated in the formation of oncogenic translocations in lymphoid tissues, including the *Myc*/*IgH* translocation in murine B cells (Boboila et al., 2012b; Simsek et al., 2011a; Simsek and Jasin, 2010). cNHEJ suppresses the formation of such chromosomal translocations (Ferguson et al., 2000). To determine the role of POLQ in chromosomal translocations, *Polq*$^{+/+}$ *and Polq*$^{-/-}$ naïve splenic B cells were stimulated in culture and assayed for the frequency of *Myc/IgH* translocations (**Figure 15A**). Notably, in the absence of *Polq* there was a 4-fold increase in translocation frequency (**Figure 15B and C**). This indicates that mammalian POLQ acts in a subset of altEJ events to suppress chromosomal

translocations. Additionally, an increase in intramolecular *IgH* rearrangements was found in B cells lacking *Polq* (**Figure 15B**). Therefore, although POLQ is involved in an altEJ pathway, it prevents rather than promotes chromosomal instability, rearrangements and the formation of *Myc*/*IgH* translocations.

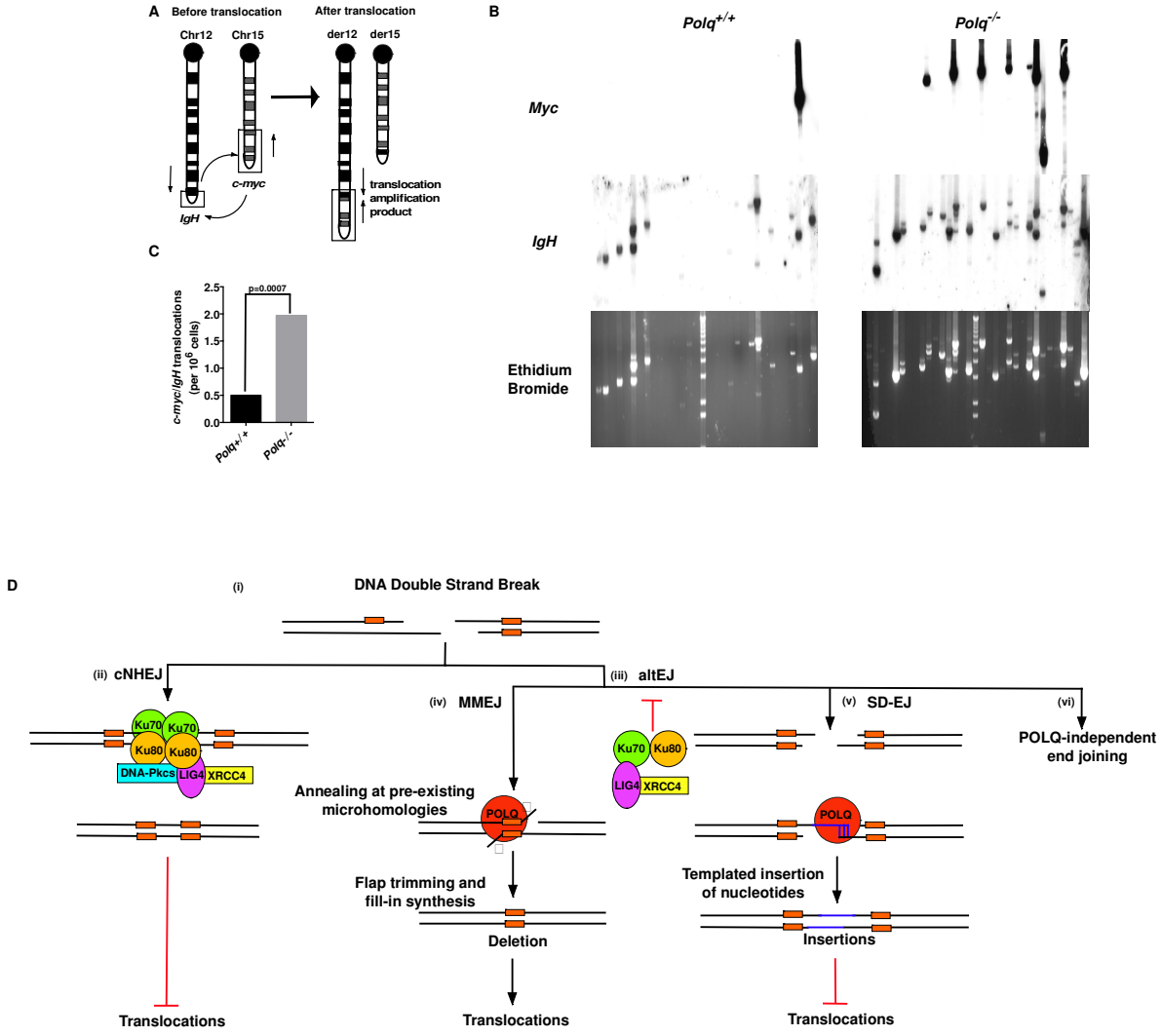


Figure 15. POLQ suppresses chromosomal translocation *in vivo*. (**A**) Representative schematic for the *Myc*/*IgH* translocation assay. PCR amplification primers are represented by black arrows. Closed circles denote centromeric locations on the chromosomes. Naïve B

cells from wild-type (WT) or *Polq*$^{-/-}$ mice were assayed for translocations after 72 h in culture. (**B**) Representative agarose gels stained with ethidium bromide and Southern blots probed with *IgH* and *Myc* probes as indicated. Each lane contains the DNA content of 1 $\times 10^5$ genomes. Three independent experiments were performed. (**C**) Frequency of translocations was plotted and p-values determined using the two-tailed Fisher's exact test. Frequencies were calculated from the total number of translocations (*Polq*$^{+/+}$: 5; *Polq*$^{-/-}$: 17) divided by the total number of genomes surveyed (9.6×10^6). (**D**) Model of end joining-mediated repair of DNA double-strand breaks (DSBs). (i) Schematic representing a DSB with existing microhomologies shown in orange. (ii) DSBs are preferentially processed by classical non-homologous end joining (cNHEJ), dependent upon Ku70–80 and Ligase4-XRCC4. This pathway is not thought to promote DNA translocations. In the absence or impairment of critical cNHEJ factors (iii) alternative end joining (altEJ) pathways are utilized. These pathways appear to be suppressed by Ku70–80 and Ligase4-XRCC4. The MMEJ pathway (iv) can orchestrate annealing of ends at pre-existing microhomologies (2–5 bp) resulting in a net deletion of genomic information. Utilization of this pathway can enhance the formation of chromosomal translocations. In the SD-EJ pathway, also known as SD-MMEJ, (v) POLQ can catalyze extension of minimally paired 3′ single-stranded DNA ends (shown in blue) to facilitate end joining and suppress the formation of chromosomal translocations. POLQ-independent alternative end joining pathways (vi) may exist but are still unclear. B cell isolation and culture and DNA isolation was performed by MJY. Sean Hensley performed PCR amplification of translocation derivative and Dr. Yunxiang Mu performed Southern Blotting.

2.4 Conclusion

POLQ suppresses hypersensitivity to direct DNA double-strand breaks

Taken together, the data presented in this chapter show that in mammalian cells, POLQ has a specific role in defense against DNA damaging agents that directly cause DNA double-strand breaks, including ionizing radiation, bleomycin, and certain topoisomerase inhibitors. Our findings indicate that POLQ participates in a novel pathway of alternative-end joining of DSBs, a process that can occur throughout the cell cycle in mammalian cells (Thompson, 2012). The minimal additional sensitization to camptothecin by olaparib in *Polq*-defective cells suggests that one function of PARP is to participate in a *Polq*-dependent altEJ pathway, which could serve as a backup pathway in homologous recombination-deficient cells. Our experiments indicate that POLQ is an important factor in DNA DSB repair in all cells, not just cells of the hematopoietic lineage. Indeed, *Polq* is broadly expressed in murine tissues (**Figure 16**).

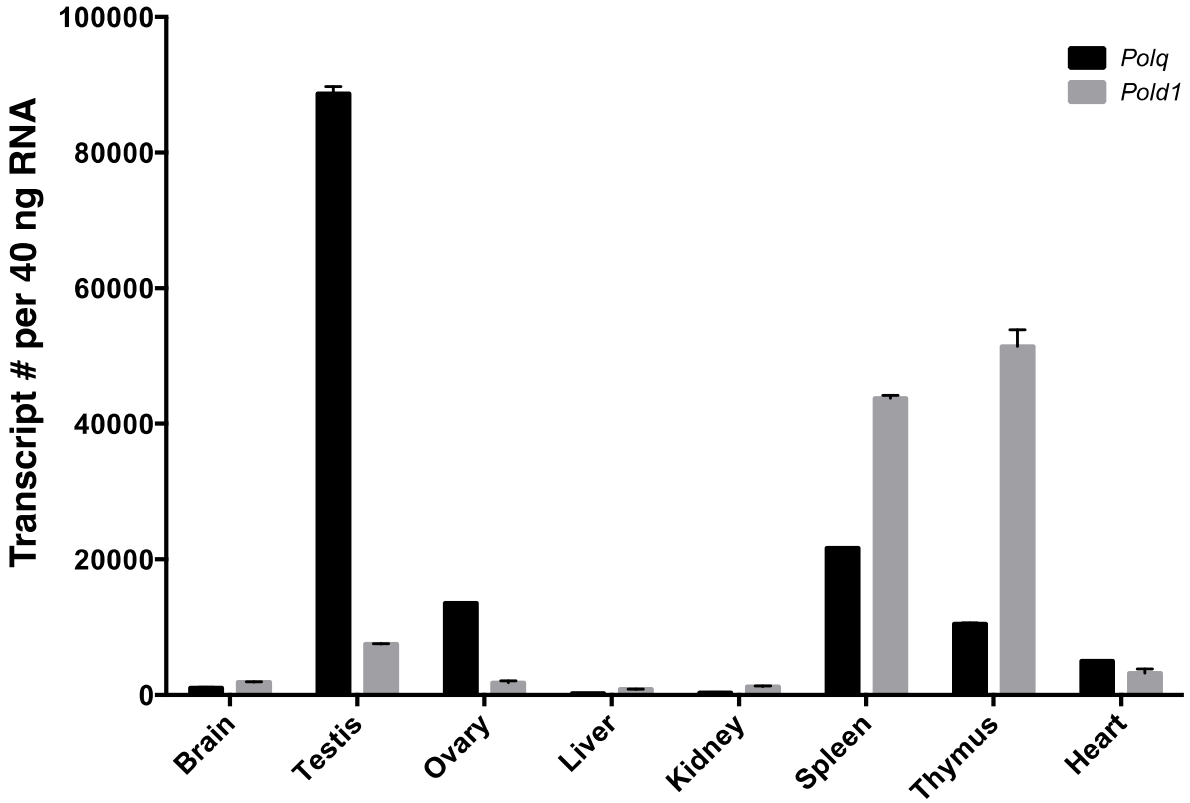


Figure 16. Absolute quantification of transcript number shows that *Polq* is broadly expressed in tissues. Total RNA was isolated from necropsied C57BL/6 mice using Trizol (Life Technologies). Transcript abundance for *Polq* and DNA polymerase delta catalytic subunit *Pold1* was determined by the absolute quantification method as described in the Material and Methods. *Gapdh* was used as an internal control to normalize samples. Tissue collection and RNA isolation were performed by MJY with assistance from Dr. Kei-ichi Takata. qPCR probe and primer sets were designed by MJY and Dr. Kei-ichi Takata. qRT-PCR was performed by the Science Park Molecular Biology Core and the data were analyzed by MJY and Dr. Kei-ichi Takata with the generous assistance of Dr. John Repass.

Mutants of POLQ homologs in *Arabidopsis* (TEBICHI), *C. elegans* (*polq-1*), and *Drosophila* (Mus308) are hypersensitive to ICL-inducing agents (Yousefzadeh and Wood, 2013), whereas *Polq*-defective mammalian cells are not appreciably hypersensitive to such agents (**Figure 5**). This difference may arise because of differences between organisms in the priority of DNA repair pathway engagement. In proliferating mammalian cells, ICLs are usually dealt with through the Fanconi anemia pathway, which produces enzymatically induced double-strand breaks that are channeled into homologous recombination repair (Kottemann and Smogorzewska, 2013). In *Drosophila* and some other organisms, an altEJ-dependent pathway may be more important for resolving ICL-associated double-strand breaks. Although *Drosophila Mus308* mutants are not hypersensitive to IR, pronounced IR sensitivity occurs in a double mutant when HR is also inactivated (Chan et al., 2010). The phenotypic consequences of POLQ-dependent altEJ repair of double-strand breaks may thus depend on the relative dominance of HR, which varies between organisms.

In this chapter I show that the DNA polymerase activity of POLQ is necessary to prevent cell death and chromosome breaks (micronuclei) caused by a double-strand break-inducing agent (**Figure 5 and 6**). Disruption of the ATPase activity in the helicase-like domain of POLQ did not, however, alter the correcting function of POLQ addition to knockout cells (**Figure 7**). A previous study with mouse cell lines suggested that disruption of the polymerase domain of the murine *Polq* gene has less severe consequences than complete disruption of *Polq* (Li et al., 2011), but the result is difficult to evaluate in the absence of quantitative measurements of expression of the disrupted form. No activity has yet been shown for the helicase-like domain, other than a DNA-dependent ATPase function (Seki et al., 2003). It is likely that an additional role remains to be discovered that is dependent on the ATPase function of POLQ.

POLQ aids DNA double-strand break repair through alternative end joining and nucleotide insertions

When double-strand breaks form in mammalian cells, a majority will be repaired through cNHEJ. However, a subset of these breaks will be handled by alternative end-joining pathways, for example, in situations where the DNA end is not compatible with processing by Ku-dependent cNHEJ, or if core components of the cNHEJ machinery are absent or unavailable (**Figure 15D**). In general, altEJ is defined as a DNA repair mechanism distinct from Ku-dependent, classically defined NHEJ (Deriano and Roth, 2013), and dependent on other factors (CtIP, MRN) that resect double-strand breaks to generate extended 3′ ssDNA tails (Bennardo et al., 2008; Lee and Lee, 2007) (**Figure 12A and 15D**). Accordingly, we showed that joining of a “pre-resected” extrachromosomal substrate (substrate with 45 nucleotide 3′-ssDNA tails) was stimulated in *Ku*-deficient cells, similar to results using chromosomal substrates (Bennardo et al., 2008). Joining of this substrate was also dependent on *Polq* (**Figure 12C**). Thus, our experiments define an altEJ subpathway in

mammalian cells that involves POLQ (termed synthesis-dependent end joining, SD-EJ or SD-MMEJ) (**Figure 15D**). Additional POLQ-independent altEJ subpathways may also be operational (**Figure 15D**). To some extent, different end-joining pathways can be been distinguished from one another by the ligase employed in the pathway, with DNA ligase IV (LIG4) suggested as essential for cNHEJ, and DNA ligase III (LIG3) for altEJ in mammalian cells (Frit et al., 2014; Simsek et al., 2011a; Simsek and Jasin, 2010). There are caveats, however. For example, some functional redundancy is apparent between LIG1 and LIG3 in altEJ (Arakawa et al., 2012; Boboila et al., 2012b; Han et al., 2014; Simsek et al., 2011b). Ligase deficiencies may thus not be the best means for distinguishing different end-joining pathways. For the altEJ subpathway under consideration here, dependence on POLQ is the best available definition.

The biochemical properties of POLQ provide a mechanistic explanation for its contribution to altEJ. POLQ has a unique ability to add nucleotides to the 3′ ends of single-stranded DNA (Hogg et al., 2012), primed by minimal pairing with other available DNA molecules (**Figure 13** and **Figure 17**). The distinct ability for POLQ to catalyze DNA synthesis from the minimally paired ends is supported by recent structural evidence that POLQ uniquely grasps the primer terminus via additional protein-DNA contacts that are not present in bacterial A-family DNA polymerases (Zahn et al., 2015). Synthesis by POLQ in this context is consistent with the unusually efficient ability of the polymerase to extend from mismatched DNA termini (Seki et al., 2004; Seki and Wood, 2008), and its tendency towards primer-template slippage (Arana et al., 2008). In future studies it will be of interest to examine the action of POLQ and DNA ligases at double strand breaks with 3′-single-stranded overhangs that closely mimic the resected ends of a DNA double-strand break. *In vivo* studies with such substrates, including those that can form hairpins in the single-stranded region, would give insight as to the preferred structures for POLQ-catalyzed extension.

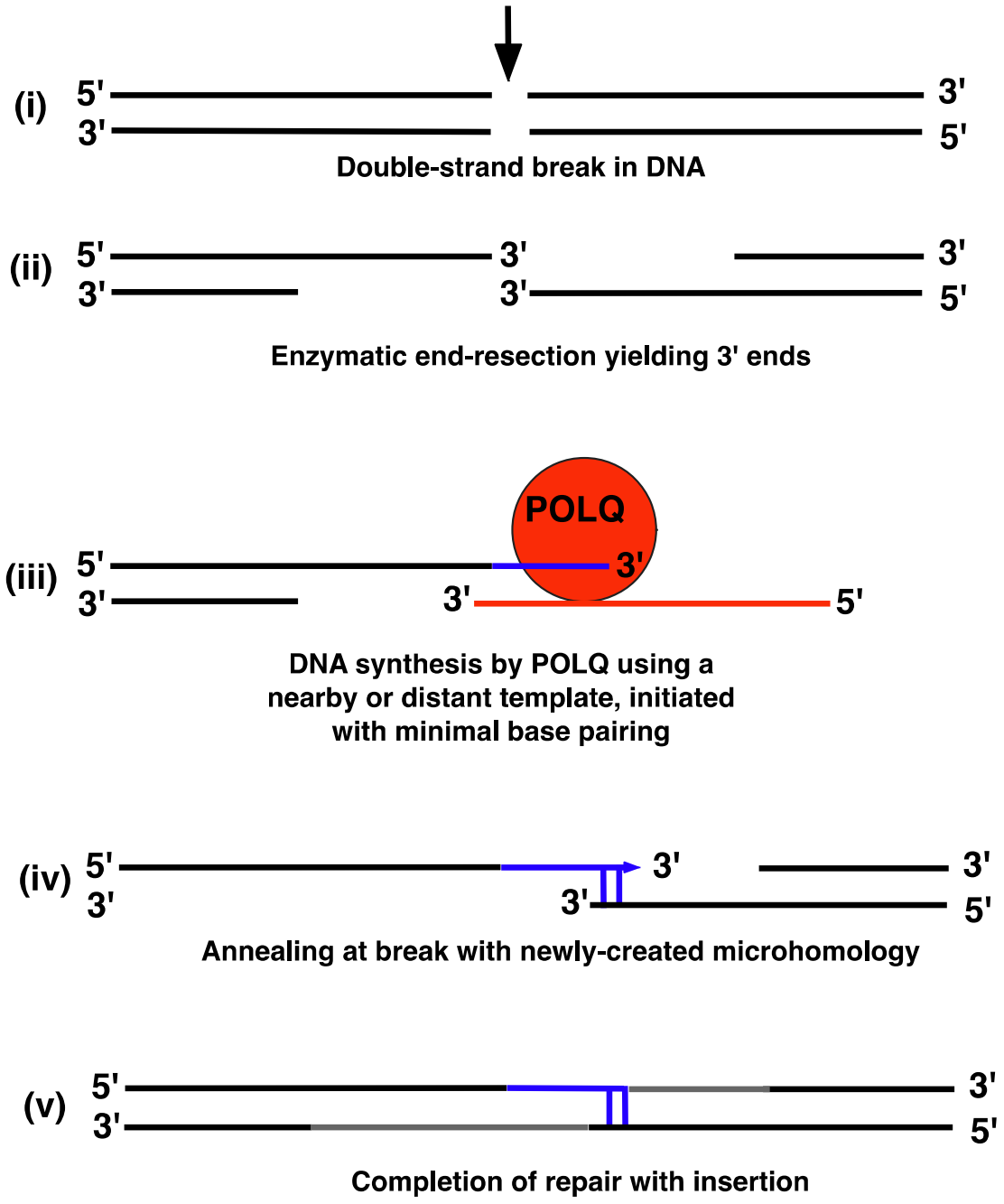


Figure 17. Mechanism of insertion formation by POLQ during double-strand break repair. After a double-strand break is formed (i), the broken ends are frequently resected enzymatically to form 3' single-stranded tails (ii). POLQ can extend a 3' end by templated

synthesis from another available DNA strand (iii). This may be the 3'-tailed end near the break, or a DNA strand available at a more distant location, or through possible "snapback synthesis" whereby the tail serves as its own intramolecular template. POLQ has the unique ability to prime DNA synthesis (blue) with minimal base pairing, sometimes slipping in the process (**Figure 14**). The newly synthesized DNA end then anneals via microhomology to the other 3' tail at the break (iv), and repair is completed (v) after further DNA synthesis (gray). This process results in an insertion of DNA sequences (blue).

A unique marker of the POLQ-dependent altEJ process is repair junctions that display templated DNA insertions. Some form of altEJ has been implicated in resolution of a subset of double-strand break intermediates in CSR, producing templated insertions (Boboila et al., 2012a). Our data support a role for POLQ in generating the CSR products with these templated insertions. These events are consistent with the templated insertions that occur during Mus308-dependent repair of directed double-strand breaks in *Drosophila* (Chan et al., 2010; White and Lambowitz, 2012) and in *C. elegans* (Koole et al., 2014). In the absence of POLQ, the lack of insertion-containing joins is observed, but the global CSR frequency is relatively unchanged (**Figure 11**). These insertions are best explained by repeated initiation of synthesis by POLQ (**Figure 14**) on template sites, ultimately leading to a joined product. An open question is whether these newly created POLQ-dependent insertions create novel microhomologies (2-5 bp) that allow for annealing of ends or the insertions produce a 3' DNA end that anneals through minimal pairing of a single base. Further analysis into the end joining properties of POLQ is warranted.

POLQ prevents the formation of Myc/IgH chromosomal translocations

In the absence of POLQ, we found a ~4-fold increase in the formation of the oncogenic translocation *Myc*/*IgH* in mice (**Figure 15C**). This increase is comparable to that seen in B cells that have lost *Tdrd3*, a regulator of R-loop formation during transcription (Yang et al., 2014) and *miRNA-155* which regulates AID and suppresses oncogenic translocations (Dorsett et al., 2008). In the absence of *Polq* there is also an apparent enhancement of rearrangement events in the *IgH* locus, consistent with the elevated level of chromosomal instability observed in cells lacking POLQ (Roerink et al., 2014).

altEJ is typically associated with frequent annealing of the DNA ends at existing microhomologies (2–5 bp) and large deletions at repair junctions (Decottignies, 2013). Since translocations commonly feature such microhomologies at their breakpoint junctions (Bunting and Nussenzweig, 2013; Weinstock et al., 2006) and occur more frequently in cNHEJ defective cells, altEJ is considered the primary mechanism by which translocations occur. Thus, a striking finding of the present work is that the formation of *Myc/IgH* translocations is suppressed when the POLQ-dependent altEJ subpathway is operational. It is possible that DNA DSBs persist for a longer time in the absence of POLQ, giving more opportunity for the formation of translocations. Alternatively, the POLQ-dependent pathway may be the most efficient at repairing a structurally distinct class of translocation-prone DNA breaks.

These studies clearly define a role for POLQ in DNA repair and provide a mechanism for template-dependent extension of DNA ends necessary to repair breaks in a subpathway of altEJ. This distinct altEJ pathway is required to prevent the formation of *Myc*/*IgH* translocations as shown by our *in vivo* experiments.

Chapter 3. Does a POLQ "signature" exist in translocations and immunoglobulin V regions.

3.1 Introduction

As just described, there is an "insertional" signature for POLQ in the immunoglobulin (*Ig*) switch regions of murine B cells. Sequencing of the switch regions showed that POLQ was required for longer insertions observed at the Sμ-Sγ1 junctions of B cells that result in IgG1+ antibodies displayed on the B cell surface (**Figure 11, Table 4**) (Yousefzadeh et al., 2014). These longer insertions feature homologies found up to 2-5 kb away from the breakpoint junction and this synthesis-dependent end joining (SD-EJ) function of POLQ is conserved in homologs found in *Drosophila* and *C. elegans* (Chan et al., 2010; Roerink et al., 2014; Yousefzadeh et al., 2014). This alternative end joining (altEJ) pathway is independent of the canonical non-homologous end joining (cNHEJ) factors Ku70 and Ku80 (**Figure 11 and 12**) (Yousefzadeh et al., 2014). POLQ-mediated SD-EJ is supported by biochemical evidence of POLQ's ability to prime synthesis from minimally-paired ends, +1/-1 slippages, and to extend from a mismatched terminus (Arana et al., 2008; Seki and Wood, 2008; Yousefzadeh et al., 2014).

This POLQ-dependent subpathway of altEJ repair of DNA double-strand breaks (DSBs) is hypothesized to promote chromosomal translocation; however the mechanism through which this occurs is still unclear (Simsek and Jasin, 2010; Zhang and Jasin, 2011). We analyzed translocations between the *Ig* heavy chain locus and the *Myc* oncogene and found a 4-fold increase in the translocation frequency in *Polq*-null B cells when compared to wild-type controls. Our findings were surprising in that we found that POLQ-dependent SD-EJ, an altEJ pathway, is actually suppressive to chromosomal translocations, which breaks with the current paradigm (**Figure 15**) (Yousefzadeh et al., 2014).

During the development of B cells, programmed induction of DNA damage is required for the diversification of immunoglobulin genes. RAG1/2 are required for initial

assembly of antibody gene segments. Further immunological diversification through class switch recombination (CSR) and somatic hypermutation (SHM) occur later on in an activation-induced cytosine deaminase (AID)-dependent fashion (Chahwan et al., 2011; Di Noia and Neuberger, 2007; Stavnezer et al., 2008). U:G mismatches, created when AID deaminates cytosines, are processed by uracil-DNA glycosylases and error prone polymerases to generate point mutations in SHM (Rajewsky, 1996). DSBs are also generated in an AID-dependent fashion that initiates CSR, causing a change in antibody isotype class without modifying antigen affinity (Chahwan et al., 2011; Di Noia and Neuberger, 2007; Stavnezer et al., 2008). The bulk of AID activity is targeted at the *Ig* locus, but AID also targets a number of transcribed genes, including tumor suppressor genes (*p53*) and oncogenes (*MYC*), which are associated with lymphomas (Gazumyan et al., 2012; Yamane et al., 2011). The DNA DSBs generated by AID are common substrates for chromosomal translocations like translocations between *IgH* and *MYC*, which are commonly observed in Burkitt lymphoma patients (Gazumyan et al., 2012; Ramiro et al., 2006).

During the repair of AID-generated DSBs, both cNHEJ and altEJ are involved but have divergent roles in the formation of chromosomal translocations. Previously altEJ, and specifically the microhomology-mediated end joining (MMEJ) pathway, was shown to promote translocations, suggesting that altEJ may be the primary pathway underlying these chromosomal aberrations in mice, while translocations appear to be dependent upon cNHEJ in humans (Boboila et al., 2012a; Ghezraoui et al., 2014; Simsek and Jasin, 2010). Our data suggests that the SD-EJ pathway that is dependent upon POLQ-catalyzed extension of 3' DNA ends serves as a genome stability mechanism that limits chromosomal translocations (Yousefzadeh et al., 2014). We hypothesize that POLQ suppresses translocations at the cost of nucleotide insertions and deletions (indels). However the impact of POLQ on translocation at the genomic sequence level is not known and would be of great value. We

hypothesize that the sequence composition of breakpoint junctions from translocations that occur when POLQ is impaired or absent are different from those that occur in the presence of POLQ.

Indels can disrupt the open reading frames of genes, but in the current context, these indels are another important antibody diversification mechanism. Nucleotide additions and deletions are found in V(D)J junctions during RAG-mediated recombination and in the variable (V) region of *Ig* genes during SHM (Helmink and Sleckman, 2012; Sale and Neuberger, 1998). The latter group of indels are termed somatic hypermutation associated (SHA) indels and can be found in the antigen-binding complementarity determining region (CDR) of Ig V regions and result from SHM (Reason and Zhou, 2006). Approximately 2-6% of peripherally-circulating human B cells feature SHA indels in their V regions (Briney et al., 2012; Wilson et al., 1998). While the frequency of these indels is low, they are crucial for antibody response to pathogens like influenza, HIV-1, and *S. pneumonia* (Krause et al., 2011; Reason and Zhou, 2006; Yu et al., 2008). Certain individuals infected with HIV-1 (termed “elite controller individuals) harbor neutralizing antibodies to HIV-1. Sequencing revealed that some of these neutralizing antibodies have nucleotide insertions in the framework regions either the heavy or light chain V regions, and functional analysis affirms that these indels are required for the neutralizing capacity of these antibodies (Klein et al., 2013). Therefore, these insertions create significant structural alterations resulting in novel antigen-antibody interactions. SHA-indels present a unique and understudied antibody diversification mechanism that is associated with SHM, but is distinct from point muations that arise during SHM. Our prior study supports a role for POLQ-dependent insertions in CSR junctions and we hypothesize that SHA-indels are POLQ-dependent as well.

3.2 Material and methods

Experimental Mice

Polq-heterozygous (*Polq*$^{+/-}$) (Shima et al., 2004) and wild-type mice were obtained from Jackson Laboratories and maintained in a C57BL/6J background. *Polq*-heterozygous mice were crossed to generate *Polq*-null mice. Mice were maintained in specific pathogen free conditions before experimental mice (2-6 mo) were transferred to a holding room that was positive for mouse parvovirus, minute virus of the mouse, and *H. pylori*. Mice were sacrificed by CO_2 euthanasia followed by cervical dislocation. Three mice per genotype were used for four independent experiments. Peyer's patches were isolated from the mice and filtered through 70 μm filters into 1X PBS. Cells were stained with the following: Propidium Iodide (PI), anti-CD19-APC, anti-CD19-PE, anti-CD95-PE-Cy7, and anti-GL7-FITC. Cells were then sorted for germinal center B cells based on the following scheme: PI^-, $CD19^{++}$, $CD95^+$, and $GL7^+$. Cells were plated into 96 well plates at 50-100 cells per well and frozen at -80℃. RNA isolation and cDNA amplification of IgK and IgH V(D)J regions were performed by the McBride laboratory. 2 μL of PCR product was TOPO TA (Life Technologies, Carlsbad, CA) cloned according to manufacturer's specifications and sequenced using M13 reverse primers at Molecular Biology Core Facility. Sequences were analyzed via the IgBLAST database (NCBI).

3.3 Results and Future Work

Immunoglobulin V region sample preparation and sequencing

In order to examine V region indels in *Polq* mice, Peyer's patches were isolated with Dr. Kevin McBride from mice and germinal center B cells were sorted by FACS. Immunoglobulin transcripts were amplified with the assistance of Sean Hensley and Monika Zelazowska and prepared for next generation sequencing analysis to determine if *Polq* status had an effect on V region indels. Before initiating next generation sequencing of immunoglobulin V regions in *Polq* mice, I wanted to ascertain population diversity of amplified transcripts. Initial sequencing of immunoglobulin genes showed numerous clonal populations, along with clonal derivations and some unique immunoglobulin genes by (**Table 6**). Following this, samples were prepared and submitted to the Next Generation Sequencing Core for analysis on the MiSeq sequencing platform (Illumina, San Diego, CA) using 300 bp paired end reading. NGS data will be analyzed by the Bioinformatics Core at the Science Park Molecular Biology Core. Our hypothesis is that based upon our findings that insertions in CSR junctions are POLQ-dependent that in the absence of *Polq* we will see a significant reduction in SHA-indels (Yousefzadeh et al., 2014). Future experiments could include the investigation of POLQ influencing antibody repertoire and responses to model antigens.

Identical Clones

genotype	sequence	clone ID	V gene	D gene	J gene	Mutations
Polq$^{+/+}$	*IgK*	K6, K17	V19-93*02	ND	J5*01	3
Polq$^{+/+}$	*IgH*	M3, M6	V1-67*01	D2-5*01	J3*01	5
Polq$^{+/+}$	*IgH*	M7, M16	V6-3*01	D1-3*01	J4*01	12
Polq$^{-/-}$	*IgK*	K6, K8	V3-4*01	ND	J5	9
Polq$^{-/-}$	*IgH*	M5, M11	V14-2*01	D3-3*01	ND	11
Polq$^{-/-}$	*IgH*	M8, M9	V7-3*01	D1-1*02	J4*01	8
Polq$^{-/-}$	*IgH*	M10, M13	V2-9*01	D2-3*01	J1*03	8

Clonal Derivation

genotype	sequence	clone ID	V gene	D gene	J gene	Mutations
Polq$^{+/+}$	*IgK*	K2	V3-12*01	ND	J2*01	10
Polq$^{+/+}$	*IgK*	K20	V3-12*01	ND	J2*01	9
Polq$^{+/+}$	*IgK*	K4	V14-111*01	ND	J4*01	10
Polq$^{+/+}$	*IgK*	K10	V14-111*01	ND	J4*01	7
Polq$^{+/+}$	*IgK*	K13	V14-111*01	ND	ND	7
Polq$^{+/+}$	*IgK*	K21	V14-111*01	ND	J5*01	8
Polq$^{+/+}$	*IgH*	M12	V7-3*01	D3-2*01	J4*01	10
Polq$^{+/+}$	*IgH*	M20	V7-3*01	D3-2*01	J4*01	11
Polq$^{+/+}$	*IgH*	M13	V1-9*01	D4-1*02	J2*01	11
Polq$^{+/+}$	*IgH*	M18	V1-9*01	D4-1*02	J2*01	12
Polq$^{-/-}$	*IgK*	K2	V3-4*01	ND	J5	12
Polq$^{-/-}$	*IgK*	K7	V3-4*01	ND	J5	12
Polq$^{-/-}$	*IgH*	M4	V7-3*01	D1-1*02	J4*01	9
Polq$^{-/-}$	*IgH*	M6	V2-9*01	D2-3*01	J1*03	12

Unique Clones

genotype	sequence	clone ID	V gene	D gene	J gene	Mutations
Polq$^{+/+}$	*IgK*	K1	V6-23*01	ND	J2*01	0
Polq$^{+/+}$	*IgK*	K5	V3-10*01	ND	J4*01	5
Polq$^{+/+}$	*IgK*	K7	V3-7*01	ND	J2*03	18
Polq$^{+/+}$	*IgH*	M10	V1-85*01	D1-1*02	J2*01	16
Polq$^{+/+}$	*IgH*	M21	V1-34*01	D2-5*01	J1*03	17
Polq$^{-/-}$	*IgK*	K3	V14-111*01	ND	J4*01	9
Polq$^{-/-}$	*IgH*	M1	V6-3*01	D4-1*02	J2*01	18
Polq$^{-/-}$	*IgH*	M5	V14-2*01	D3-3*01	ND	11
Polq$^{-/-}$	*IgH*	M17	V14-3*01	D2-12*01	J3*01	6

Table 6. Sequencing of *IgH* and *IgK* sequences from *Polq* mice. cDNA amplicons were TOPO TA cloned and sequenced. Sequences were analyzed by IgBLAST. Individual V, D, J recombinants are denoted and amino acid mutations are enumerated. Clonal sequences with identical VDJ gene segments and mutations, their clonal derivatives featuring novel mutations (bracketed into groups), and unique sequences are stratified. Germinal Center B

cells were isolated by Dr. Kevin McBride and MJY. Immunoglobulin transcript amplification was performed by Sean Hensley, Monika Zelazowska, and MJY. MJY cloned the cDNA amplicons and analyzed the sequencing results generated by the Science Park Molecular Biology Core.

Translocation analysis is *Polq/Tp53* mice

In the absence of *Polq*, *Myc/IgH* translocation frequency increases nearly four-fold relative to the wild-type control (**Figure 15)** (Yousefzadeh et al., 2014). However, in stimulated B cells from wild-type mice, the translocation frequency is quite low (~1 translocation per 500,000 cells), making translocation analysis reagent-intensive (**Figure 15)** (Yousefzadeh et al., 2014). Therefore I am crossing of *Polq* mice to *Tp53* mice to enhance the translocation frequency. *Tp53*-null mice exhibit a 5 to 10-fold increase in the frequency of *Myc/IgH* translocations compared to wild-type mice (Ramiro et al., 2006). Our hypothesis is that POLQ prevents the formation of chromosomal translocations but it comes at the cost of creating indels. Thus it is important to learn the locations and compositions (insertions, microhomologies, etc.) of translocation junctions that are dependent upon POLQ. C57BL/6J-*Polq*- and FVB-*Tp53*-null mice are being crossed to generate *Polq*$^{-/-}$ *Tp53*$^{-/-}$ mice and will be available for use soon. These mice will be used for analysis of translocation breakpoints by NGS. Future studies could spring from this initial analysis such as the investigation of POLQ's impact on other common translocations.

Chapter 4: The effect of POLQ suppression on radiosensitivity of breast cancer cells.

4.1 Introduction

For many breast cancer patients, surgical tumor resection is followed by treatment with ionizing radiation (IR). A salient need in breast cancer therapeutics is to enhance the effectiveness of IR. “Triple-negative” breast cancer (TNBC), a particular subset of breast cancers, are refractory to most conventional therapies (André and Zielinski, 2012), including IR, and pose a major clinical challenge (Lehmann et al., 2011). As previously reported, cells lacking POLQ are notably hypersensitive to IR and the radiomimetic bleomycin (Goff et al., 2009; Higgins et al., 2010b; Yousefzadeh et al., 2014). Our lab and others demonstrated that abnormally high levels of POLQ in breast cancer patients correlates with poorer outcomes (Begg, 2010; Higgins et al., 2010a; Lemée et al., 2010). Notably POLQ is a component in the Genomic Grade Index (GGI) signature and “76-gene signature” set which are used to assess tumor grade and prognosis (Sotiriou et al., 2006; Wang et al., 2005).

While multiple studies have delved into the correlation between *POLQ* expression and breast cancer outcomes, very little was known about *POLQ* single nucleotide polymorphisms (SNPs) and breast cancer until recently. A SNP of *POLQ* (c.-1060A>G) was shown to be significantly more frequent in hereditary breast cancer patients than in sporadic breast cancer patients or control individuals (Brandalize et al., 2014). This SNP is located in the promoter region of *POLQ* but it is not known how this SNP affects *POLQ* expression or how it modifies the disease in these hereditary breast cancers.

The mounting evidence supporting a role for POLQ in breast cancer, makes it an ideal new therapeutic target. Targeting DNA replication as a therapeutic intervention is a time tested and clinically used principle. For example chemotherapeutics like 5-fluorouracil and hydroxyurea, and nucleoside analogs such as gemcitabine and cytarabine are commonly used in the treatment of cancer patients (Kelley, 2011). Since POLQ is the only

DNA polymerase to be significantly overexpressed in breast cancer, is involved in replication timing, and is not essential for cell viability (Fernandez-Vidal et al., 2014; Lemée et al., 2010; Yousefzadeh and Wood, 2013). POLQ a logical choice and appealing target for therapeutic development.

4.2 Material and methods

Cell lines

The following cell lines were purchased from American Type Culture Collection: MCF7, MCF10, HCC1569, HCC1599, HCC1806, BT-474, MDA-MB-361, and MDA-MB-436. Cells were STR fingerprinted, checked for mycoplasma contamination, and cultured in their appropriate media (Neve et al., 2006).

RNA isolation and qPCR analysis

Total RNA was isolated from the each of the breast cancer cell lines. RNA quality and purity was analyzed using a RNA 6000 Nano kit (Agilent Technologies, Santa Clara, CA). 1 µg of total RNA was used to generate cDNA using the High Capacity cDNA RT kit (Life Technologies). qPCR analysis was performed in triplicate using the ABI Prism 7900 HT thermocycler and the following Taqman Probe set or primer set with iTAQ SYBR Green Supermix with ROX (Bio-Rad, Hercules, CA): (*POLQ*) Qexon FWD 5′-TGCCTTTCAAAAGTGCCCGGAAGGC3′, Qexon REV 5′-TGCCAGTCACCCANATAGTTCNCAT-3′ (where N represents a dSpacer); (*POLD1*) HsPOLD1_FWD 5′-CGACCTTCCGTACCTCATCTCT-3′, HsPOLD1_R 5′-ACACGGCCCAGGAAAGG-3′, probe 5′-CCCTCAAGGTACAAACAT-3′; (*POLN*) HsPOLN_FWD 5’-ACTGATGGTTCCACCCAGCTA-3’, HsPOLN_REV 5’-GCGTTTTACTAACACCACAATTCCT-3’, probe 5’-ACCAGACCCCCGTTTCT-3’; *HELQ* (Hs00431095_m1); *REV3L* (Hs01076848_m1); *REV7* (Hs01057448_m1); *EXO1* (Hs01116195_m1); *MCM2* (Hs01091564_m1); *RAD51C* (Hs00427442_m1). Data were analyzed using the ΔΔCt method. *GAPDH* was used as an internal control in all experiments and relative quantification analysis was performed with MCF10 cells serving as the non-tumorigenic control.

Transfections and Transductions

293T cells (kindly provided by Dr. Christopher Bakkenist, University of Pittsburgh Medical School) were plated at 1.5 x10^5 cells per well in six-well plates and transfected the following day with 5 µg of either SIGMA Mission shRNA non-targeting control (shCTRL); or five different shRNAs against POLQ (shQ46, shQ47, shQ48, shQ49, shQ50) using jetPRIME (Polyplus, New York, NY) according to manufacturer's specifications. 48 h after transfection, cells were harvested for RNA isolation (RNeasy, Valencia, CA) or for immunoblotting. To generate lentivirus used for transduction, 2.5 ×10^6 293T cells were plated and cotransfected the following day with psPAX2 (6 µg), pMD2G (6 µg), and (12 µg) SIGMA Mission shRNA vector (shCTRL, shQ46, shQ50) using jetPRIME. One day prior to transduction, HCC1806 cells were seeded into a 10 cm dish at 1.5 ×10^5 cells with 15 mL complete media. 48 h post-transfection virus-containing media was harvested, filtered through a 0.45 µm syringe filter and used to replace the media on the plated HCC1806 cells with 15 µL of 4 µg/mL polybrene (1 µL of polybrene stock per mL of complete media). Cells were incubated in the virus-containing media for 24 h before being split into T-75 flasks and grown to 80% confluence before undergoing three weeks of puromycin selection (2.5 µg/mL). Following selection, pure clones were isolated and cultured with complete media containing puromycin (1 µg/mL).

shRNA sequences

shCTRL (SHC002) 5'-CCGGCAACAAGATGAAGAGCACCAACTC-3'

shQ46 (TRCN0000290546) 5'-CCGGCCTTCAATCTTGCTTGCGAAACTCGAGTTTCGCAAGCAAGATTGAAGGTTTTTG-3' (anneals to *POLQ* cDNA sequence 5840:5861)

shQ47 (TRCN0000290547) 5'-CCGGCGGGCCTCTTTAGATATAAATCTCGAGATTTATATCTAAAGAGGCCCGTTTTTG-3' (anneals to *POLQ* cDNA sequence 2967:2991)

shQ48 (TRCN0000290548) 5'-CCGGCGTCGTCTCATTCAAGTGTTACTCGAGTAACACTTGAATGAGACGACGTTTTTG-3' (anneals to *POLQ* cDNA sequence 7036:7056)

shQ49 (TRCN0000290549) 5'-CCGGCCCTGTTACATTCTAGTACATCTCGAGATGTACTAGAATGTAACAGGGTTTTTG-3' (anneals to *POLQ* cDNA sequence 2705:2725)

shQ50 (TRCN0000290550) 5'-CCGGGCTGACCAAGATTTGCTATATCTCGAGATATAGCAAATCTTGGTCAGCTTTTTG-3' (anneals to *POLQ* cDNA sequence 699:719)

4.3 Results and future work

To evaluate our hypothesis that suppression of POLQ in breast cancer cells confers radiosensitivity, we chose a number of breast cancer cell lines of different statuses (p53, ER, PR). We are particularly interested in triple-negative breast cancers (TNBC), which, through bioinformatics analyses, have been shown to overexpress *POLQ* and are quite refractory to treatment. Relevant cells lines from the Basal A (MCF10, MDA-MB-436), B (HCC1569, HCC1806, HCC1599), and luminal (MCF7, BT-474, MDA-MB361) subtypes were obtained from ATCC and tested to ensure no mycoplasma contamination. The cell lines were then analyzed for expression of *POLQ* and other DNA replication and repair genes (**Figure 18A**) with similar expression patterns to *POLQ* as previously determined through microarray analysis of breast tumor subtypes.

In the various breast cancer cell lines *POLQ* was expressed at 8 to 114-fold higher levels than in the non-tumorigenic MCF10 control cell line (**Figure 18B**), whereas related gene family members *POLN* and *HELQ* show very little variability in expression amongst all the cell lines (**Figure 18C**). Other DNA polymerases are expressed at comparable levels between the cell lines (**Figure 18D and E**). Thus, the high degree of *POLQ* expression in multiple breast cancer cell lines, including TNBC cell lines, is unique to this DNA polymerase. These results are consistent with a previous study that also found that POLQ was the only DNA polymerase significantly overexpressed in multiple breast cancer tumors (Lemée et al., 2010). However, there was a broad range of expression patterns amongst different DNA repair genes (*EXO1*, *MCM2*, and *RAD51C*) (**Figure 18F**).

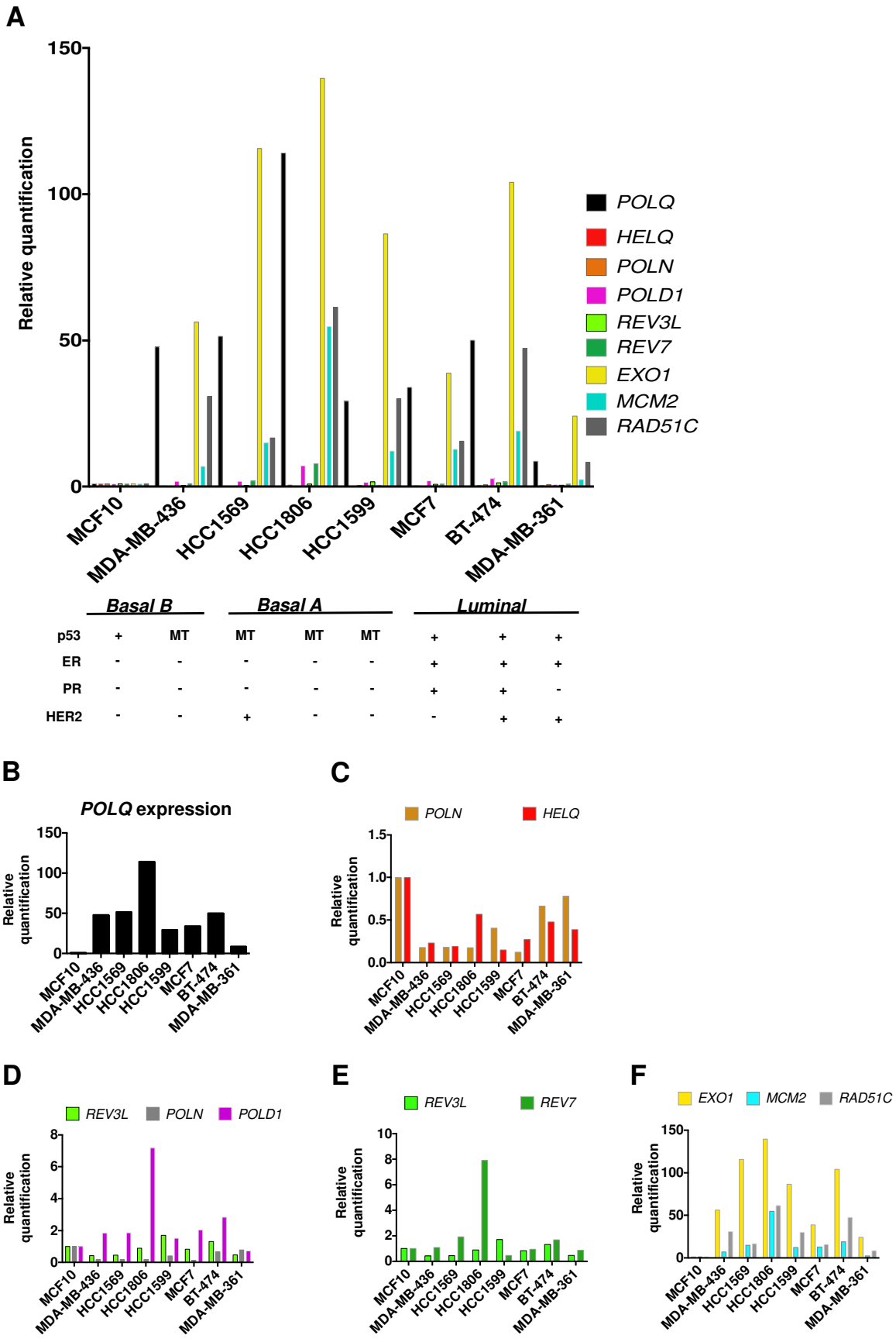

A
Relative quantification
150
100
50
0
POLQ
HELQ
POLN
POLD1
REV3L
REV7
EXO1
MCM2
RAD51C
MCF10
MDA-MB-436
HCC1569
HCC1806
HCC1599
MCF7
BT-474
MDA-MB-361
Basal B
Basal A
Luminal
p53 + MT MT MT MT + + +
ER - - - - - + + +
PR - - - - - + + -
HER2 - - + - - - + +
B
POLQ expression
Relative quantification
C
POLN
HELQ
D
REV3L
POLN
POLD1
E
REV3L
REV7
F
EXO1
MCM2
RAD51C

Figure 18. Survey of DNA polymerase and repair gene expression in breast cancer cells. Expression of the indicated genes was quantified by qPCR in multiple breast cancer cells. Relative quantification of gene expression was determined using Taqman probes, data was analyzed using the ΔΔCt method and *GAPDH* expression was used as an internal control. Abbreviations: MT (mutated), ER (presence of estrogen receptor), PR (presence of progesterone receptor), HER2 (amplification of *HER2/neu*).

Baseline *POLQ* expression levels have now been established for different breast cancer cell lines and breast cancer subtypes. These studies provide insight for selecting cell lines that will be candidates for POLQ-silencing. Initial analysis of shRNAs directed at POLQ was performed in 293T cells and RNAi-mediated suppression was measured by qPCR analysis (**Figure 19A**). Two individual POLQ shRNAs (shQ46 and shQ50) knockdown POLQ by ~50-60% alone, but did not act synergistically or additively when combined (**Figure 19B**).

Efficient shRNA suppression of POLQ is needed in order to test radiosensitivity by clonogenic assays. Using this method, we wish to test our hypothesis that cell lines with elevated levels of *POLQ* will exhibit greater radiosensitivity than cells with reduced *POLQ* levels due to shRNA treatment. As a proof of principle experiment to test this hypothesis we chose to test the HCC1806 cell line, as it had the highest relative level of *POLQ* expression when compared to the MCF10 non-tumorigenic control cell line. Stable knockdown and relevant control cell lines were generated after lentiviral transduction of shRNA constructs and puromycin selection before testing for *POLQ* suppression. *POLQ* transcript levels were reduced by ~50% in HCC1806 POLQ knockdown cells transduced with either shQ46 or shQ50 (**Figure 19C**). Clonogenic survival assays of irradiated cells are currently in progress to measure radiosensitivity upon suppression of POLQ by shRNA. We may fail to observe enhanced radiosensitivity in breast cancer cell lines if the shRNA-mediated suppression of POLQ fails to adequately suppress POLQ.

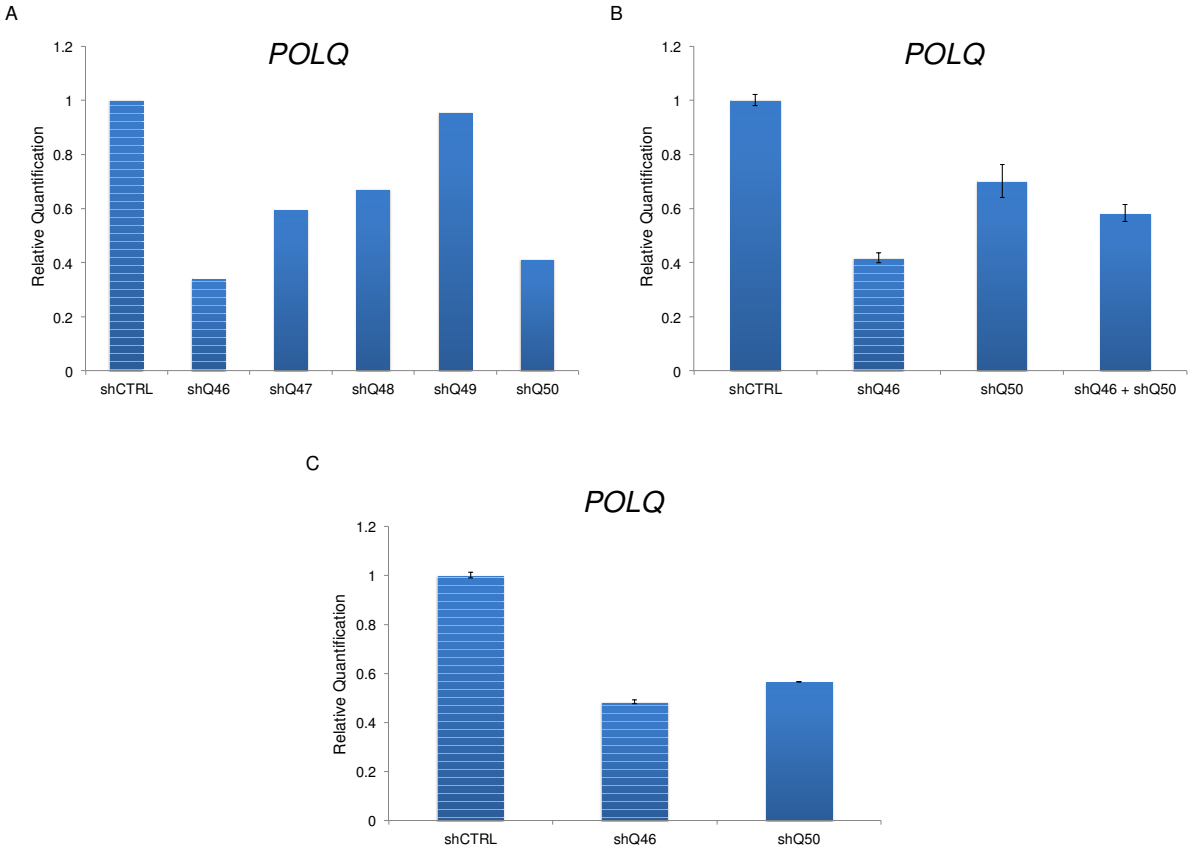


Figure 19. shRNA knockdown of POLQ in cancer cell lines. 293T cells were transfected with either shCTRL non-targeting control shRNA or an shRNA targeted against POLQ (shQ46-50). RNA was harvested from the cells 48 h post transfection and analyzed for *POLQ* expression by qPCR. qPCR analysis was performed in duplicate **(A)** or triplicate **(B)** for the indicated experiments. (**C**) HCC1806 cells stably knocked down for POLQ by shRNA were surveyed for *POLQ* expression in a similar manner. *GAPDH* was used as an internal control and ΔΔCt analysis was performed.

An alternative approach to combat this problem would be to employ CRISPR-mediated deletion of *POLQ* from these cell lines to generate multiple POLQ knockout breast cancer cell lines. Small molecule inhibitors of POLQ arising from an ongoing screen of

POLQ inhibitors being performed in conjunction with the Texas Screening Alliance for Cancer Therapeutics and Texas Institute for Drug and Diagnostic Discovery, will be screened against the breast cancer cell lines and *Polq* knockout mouse embryonic fibroblast (MEF) lines. If POLQ is the rate limiting factor for altEJ, then we hypothesize that small molecule inhibitors of POLQ will further radiosensitize breast cancer cells with aberrantly high levels of *POLQ* (responders) while having a smaller effect on radiosensitivity in those cells with lower levels of *POLQ* (non-responders). Specificity of the small molecule inhibitors will also be verified by looking at the enhancement of radiosensitization of *Polq* MEFs, where *Polq*-proficient MEFs should only show a synergistic increase in sensitivity to ionizing radiation while *Polq*-deficient MEFs should have no further increase in radiosensitivity above their already elevated baseline. A validated compound could then be further analyzed to understand its interaction with, and mechanism of, POLQ inhibition (catalytic, allosteric, etc.). This compound could then be used for *in vitro* studies in *Polq*-deficient cell lines to determine if its effects are POLQ-specific. Much work remains to be done but a small molecule inhibitor of POLQ could be of great clinical value in breast cancer therapeutics by enhancing radiosensitivity, particularly in TNBC patients, whose tumors already express aberrantly high levels of *POLQ* (unpublished data, C. Marcelo Aldaz and Richard Wood), and are quite refractory to most treatments (André and Zielinski, 2012).

Chapter 5. A-family DNA polymerases and the repair of x-ray induced DNA damage.

5.1 Introduction

The human genome is under constant attack from both endogenous and exogenous sources of damage. About 2% of our genome is committed to maintaining genomic stability. There are more than 175 proteins involved in DNA repair (Friedberg et al., 2006) (http://sciencepark.mdanderson.org/labs/wood/dna_repair_genes.html). These proteins are critical in the repair or tolerance of lesions that result from many sources such as endogenously generated reactive oxygen species or exogenous damage caused by ultraviolet radiation. The "DNA damage response" is an a collective term for the various molecular pathways including DNA repair and replication, and cellular checkpoints that occur after DNA damage. The synergistic activities of these different processes allow cells to detect DNA damage, arrest cells at given points in the cell cycle and potentially initiate destruction of the cell if the damage is too great.

Various repair pathways can repair the damage but despite the prominence of these repair mechanisms, lesions can escape the repair process. These lesions create roadblocks for the replicative polymerases (α,δ,ε), which catalyze DNA synthesis during S phase of the cell cycle (Lange et al., 2011). However, the replicative polymerases cannot bypass the damaged DNA, thereby creating stalled replication forks that can lead to the formation of double strand breaks. Replication through the damaged DNA is achieved by a group of enzymes known as translesion synthesis (TLS) polymerases.

The mammalian genome encodes seven polymerases capable of performing TLS *in vitro*: PrimPol, pol ζ, pol η, pol ι, pol κ, pol ν, and pol θ. Two models currently exist for TLS: the one-step model in which a TLS polymerase inserts a nucleotide across from the damaged template and catalyzes further extension along the template, or the two-step model, in which one polymerase inserts across the damaged template and falls off, followed

by another TLS polymerase extending past the damage (Lange et al., 2011). The presence of these specialized polymerases suggests that bypass of the DNA lesions is an important component of DNA damage tolerance.

Human DNA polymerase ν (POLN) is an A-family DNA polymerase related to *Escherichia coli* polymerase I. POLN was identified by searching for sequence similarity to the *mus308 Drosophila* DNA polymerase (**Figure 1**) (Marini et al., 2003), which was identified in a genetic screen for genes that granted resistance to DNA damaging agents (Harris et al., 1996). The human *POLN* gene is 160 kb long and encompasses 24 exons with conservation of C-terminal exons 14-24, producing a 4kb transcript and 900 amino acid protein. Many inactive alternative splicing variants were found, as a result of exon skipping or alternative exon inclusion, in cell lines (Marini et al., 2003). POLN is broadly conserved amongst the vertebrate lineage. POLN possesses DNA polymerase activity *in vitro*, including low fidelity DNA synthesis, as well as the ability to accurately and efficiently bypass thymine glycol DNA adducts and specific protein-DNA crosslinks (Takata et al., 2006; Yamanaka et al., 2010). By way of northern blot analysis, full-length POLN has been shown to be present in both human and mouse testes (Marini et al., 2003). Due to its high expression in testis, POLN may serve as a recombinase for spermatogenesis, which occurs during meiosis. POLN may not be limited to the testis. Transcripts have also been found in both human testis (833K) and hematopoietic cell lines from human (K562) and mouse (L1210) cell lines (Marini et al., 2003). Previously, POLN was reported to play a role in DNA interstrand crosslink (ICL) repair in human cells (Moldovan et al., 2010; Zietlow et al., 2009), and a role for POLN in homology-directed repair was proposed in chicken DT40 cells (Kohzaki et al., 2010). Further analysis is required to determine the *in vivo* role of POLN in mammals.

5.2 Material and methods

Mouse tissue RNA isolation

Wild-type C57BL/6 mice were sacrificed via carbon dioxide euthanasia and immediately necropsied in order to isolate their organs. The organs were placed into RNA Later buffer (Ambion) on ice and taken to the laboratory for RNA isolation. The tissues were flash frozen with liquid nitrogen and ground by mortar and pestle into a fine powder. The powder was placed into Trizol (Invitrogen) and homogenized using a 16-gauge needle to lyse the cells. Chloroform extraction (Sigma-Aldrich) was performed to purify the RNA and concentration was measured by absorbance at 260 nm using a Nanodrop spectrometer. RNA quality was ensured by first measuring the 260/280 nm ratio and then running 5 μg of total RNA on a 1% denaturing agarose gel. After ethidium bromide (Sigma-Aldrich) staining and subsequent destaining, RNA quality was measure by the absence or presence of the 18S and 28S ribosomal RNA bands visualized by UV transillumination using an Alpha Imager system.

Absolute quantification of Poln expression

1 μg of total RNA was used to generate cDNA using the High Capacity cDNA RT kit (Life Technologies). qPCR analysis was performed in triplicate using the ABI Prism 7900 HT thermocycler and the following Taqman Probe set or primer set with iTAQ SYBR Green Supermix with ROX (Bio-Rad, Hercules, CA): (*Poln*) FWD 5'-GATGTACAAGGATGGTTCCACACA-3', REV 5'-GCCTACATGGCTTTTTAGTAACACTACAAT-3', probe 5'-AAAGCCTCCTTGGCGCTCA-3'; (*Pold1*) FWD 5'-TTTGACCTCCCATACCTCATCTCT-3', 5'-ACGCGGCCCAGGAAAG-3', probe 5'-AAGCGGTCCACCTTTAG-3' For absolute quantification, titration of DEST17/MmPolN and pET/MmPold1 plasmids were used to generate standard curves. Transcript abundance was determined by extrapolation from linear regression analysis of

best-fit lines from titration experiments. *Gapdh* was used as an internal control in all experiments.

B cell culture and CSR analysis

B cells were isolated from mouse spleens (n = 6 per genotype) and stimulated for class-switching, in culture, for 72 h as described in Chapter 2 by Dr. Kevin McBride, Dr. Kei-ichi Takata, and Shelley Reh. Dr. Takata prepared genomic DNA from stimulated B cells using the DNeasy kit rather than as described in Chapter 2, but was not able to PCR amplify Sμ-Sγ1 CSR junctions. I repurified the DNA by phenol chloroform extraction and PCR amplified Sμ-Sγ1 CSR junctions were amplified as specified in Chapter 2. The resulting product was TOPO TA cloned and transformed into Top10 *E. coli* cells (Life Technologies, Carlsbad, CA). Plasmids were purified and sent for sequencing using M13 FWD and REV primers in addition to the primers used for PCR amplification. 50 clones for each group were analyzed for mutations, deletions, insertions, and sequence overlaps at the junction, and both 30 nt upstream and downstream of the junction. p-values were determined by using the two-tailed Fisher's exact test.

Generation and complementation of Poln$^{-/-}$ Polq$^{-/-}$ MEFs

Polq-heterozygous (*Polq*$^{+/-}$) mice (Shima et al., 2004) were obtained from Jackson Laboratories and were crossed to generate *Polq*-null mice. Mice were maintained in a C57BL/6J background before being crossed with C57BL/6J *Poln*$^{-/-}$ mice (developed by Kei-ichi Takata). Double heterozygous mice (*Poln*$^{+/-}$ *Polq*$^{+/-}$) were crossed to yield double knockout mice (*Poln*$^{-/-}$ *Polq*$^{-/-}$), which were used for timed pregnancies in order to generate MEFs. Isogenic primary MEFs were harvested from 13.5 d pregnant females and cultured in a 2% O_2 atmosphere. MEFs were then transfected with 1 μg of pSV-Tag (Lange et al., 2012;

Sobol et al., 1996; Yousefzadeh et al., 2014) and grown in atmospheric (~20%) oxygen for six population doublings to allow for immortalization.

DNA damage sensitivity assays

Complemented MEF lines were plated in triplicate into white 96 well plates at 1250 cells per well and grown overnight using complete media containing puromycin (1 µg/mL). The following day, cells were cultured with complete media containing the indicated amounts of bleomycin (dissolved in 150 mM NaCl) or mitomycin c (dissolved in ethanol). After 24 h of bleomycin treatment, the medium was changed and cells were allowed to recover for 72 h before cellular viability was measured using the ATPlite 1Step kit (Perkin Elmer, Waltham, MA) using a Biotek plate reader. Cells treated with mitomycin c were incubated in drug for 48 h before cellular viability was measured. Experiments were repeated three times. Both bleomycin and mitomycin c were purchased from Sigma-Aldrich (St. Louis, MO).

POLQ gain of function

293T cells were plated at 1.5 x10^5 cells per well in six-well plates and transfected the following day with 2 µg of either pCDH/POLQ (Q), pCDH/POLQ-D2330A,Y2331A (P), or empty vector control (E). The next day, cells were plated in triplicate into white 96 well plates at 1250 cells per well and grown overnight using complete media. At 48 h post-transfection, cells were cultured with complete media containing the indicated amounts of bleomycin (dissolved in 150 mM NaCl) for 72 h and cellular viability was measured using the ATPlite 1Step kit (Perkin Elmer, Waltham, MA) using a Biotek plate reader. Experiments were repeated three times.

Immunoblotting

For immunoblots, cells were resuspended in 200 µL of 2X SDS loading buffer (4% SDS, 0.2% bromophenol blue, 20% glycerol, 100 mM Tris HCl pH 6.8, 12% 2-mercaptoethanol) and heated at 95℃ for 5 min. 20 µL of extract was separated on a 4-20% polyacrylamide gel, transferred to PVDF membrane, blocked, and blotted with anti-αTubulin (Abcam, Cambridge, UK) ab4074, 1:10,000), anti-FLAG (Sigma F7425, 1:5,000), or anti-POLQ (MDACC POLQ20, 1:250) antibodies and corresponding secondary antibodies (Sigma A0168, A0545; 1:10,000), and visualized with ECL reagent (Pierce, Rockford, IL).

2X SDS Loading Buffer

20% Glycerol

4% SDS

0.2% Bromophenol Blue

12% 2-Mercaptoethanol

100 mM Tris-HCl pH 6.8

5.3 Results

Poln is highly-expressed in testis

To obtain a more definitive and precise picture of *Poln* expression, Poln transcript levels were measured using qPCR and absolute quantification (AQ) of *Poln* in various mouse tissues. Plasmids were obtained for *Pold1* and *Poln*. These plasmids were used in respective ABI Taqman assays to create a standard curve of C_T values to be able to determine absolute quantification of transcript level and copy number for *Poln* and *Pold1*. qPCR analysis showed that the highest levels of *Poln* transcripts were found in the testis (**Figure 20**), correlating with prior Northern blot data that suggested *Poln* expression is only easily detectable in the testis (Marini et al., 2003). Thus it appears amongst all the mouse tissues surveyed, POLN is most highly expressed in the testis.

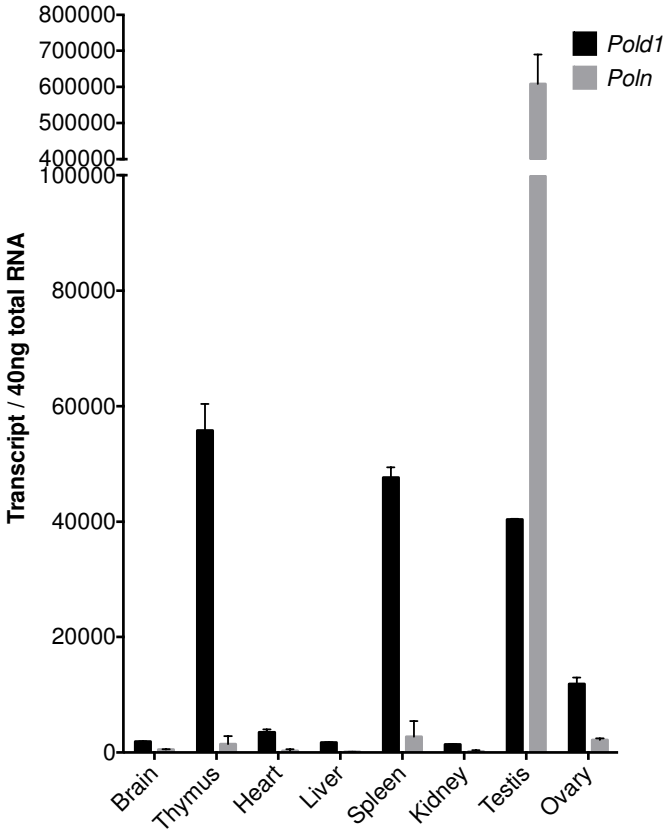


Figure 20. Absolute quantification of transcript number shows that *Poln* is highly expressed in testis. Total RNA was isolated from necropsied C57BL/6 mice by Trizol (Life Technologies). Transcript abundance for *Poln* and DNA polymerase delta catalytic subunit *Pold1* was determined by absolute quantification method as described in the Material and Methods. *Gapdh* was used as an internal control to normalize samples. Dr. Kei-ichi Takata and MJY isolated tissues, purified RNA and analyzed qPCR data generated by the Science Park Molecular Biology Core.

POLN does not operate in the POLQ-dependent altEJ pathway

Within the past few years Dr. Kei-ichi Takata of the Wood Laboratory has generated constitutive *Poln* knockout mice (data unpublished). Phenotyping of the mice is still ongoing and includes everything from aging studies to analysis of cells derived from the mice. Due to POLN being related to the POLQ gene family, we hypothesized that POLN could play a role in synthesis-dependent end joining pathway in the repair of DNA DSBs as POLQ does (Yousefzadeh et al., 2014). To test this hypothesis we performed class switch recombination analysis on naïve B cells isolated from *Poln*-deficient and -proficient mice. The frequency of $IgG1^+$ B cells in mice lacking *Poln* were comparable to levels found in the wild-type mice, an observation similar to that observed in the *Polq* mice (unpublished data, Kei-ichi Takata and Richard Wood) (Yousefzadeh et al., 2014).

In an effort to bias this end joining-dependent pathway to further utilize alternative end joining (altEJ), B cells were treated with NU7026, a DNA-PKcs inhibitor that inhibits classical non-homologous end joining (cNHEJ). While inhibition of DNA-PKcs does increase the frequency of switch as previously noted (Callen et al., 2009), there was no discernable difference in switching between genotypes (unpublished data, Kei-ichi Takata and Richard Wood). In my analysis of CSR junctions, *Poln* status, unlike *Polq* status, has no bearing on

the long insertions into CSR junctions that are thought to be altEJ-dependent (**Figure 21**) (Yousefzadeh et al., 2014). Longer insertions that are homologous to nearby sequences are still found in the absence of POLN (**Figure 21**, **Table 7**). Although POLQ may still play a role in DNA repair, it does not appear to function in the SD-EJ pathway of DNA DSB repair.

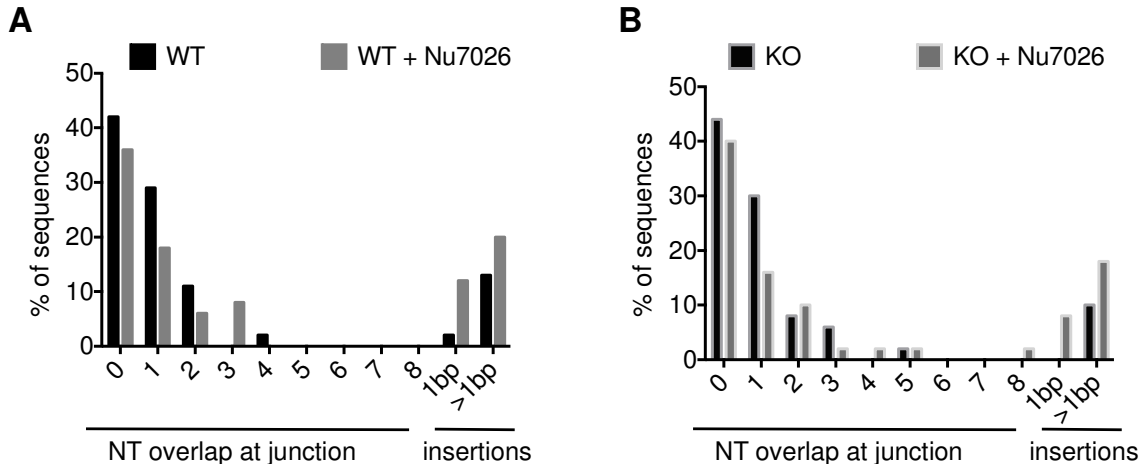


Figure 21. Analysis of CSR junctions in *Poln* mice. Isolated wild-type (WT) and *Poln*$^{-/-}$ (KO) naïve splenic B cells were stimulated for CSR and either mock-treated or treated with NU7026. Genomic DNA isolated from B cells of wild-type (**A**) and *Poln* KO (**B**) mice was amplified by PCR and 100 Sμ-Sγ1 junctions from each group were sequenced and analyzed for overlaps and insertions at breakpoints.

	Sμ (J00440)	Insertion	Sγ (D78344)		Insert Homology
	WT				
5025:5054	GAGCCCTAGTAAGCGAGGCTCTAAAAAGCA	CAGCTGAGC	CAGGTTAGAATGAAGGATGGGCATCCCGGG	2909:2938	Sγ3176:3182
5147:5176	CCAAAATGAAGTAGACTGTAATGAACTGGA	TC	TAAGCAGGGACAGGTGGAAGTGTGGAGACC	8197:8226	
5165:5194	TAATGAACTGGAATGAGCTGGGCCGCTAAG	AG	AAGTACTTGTAGAGGAACAGGGGCAGGTTA	2886:2915	
5377:5406	AGATGGGGGTGGAGATGGGGTGAGCTGAGC	TGGGCT	GAGCTAGACTGAGCTGAGCTAGGGTGAGCC	7146:7174	
5290:5319	GTTCTGG--TGAGCTGAGCTGGGGTGAGCT	GAGCTGAGCTGAGCTGGGG	TG-AGCTG-GGGTAAGCAGGGACAGGTGGA	8185:8214	Sμ5452:5470
5374:5403	AGCTGAGATGGGGTGAGATGGGGTGAGCTG	GTGGAAA	GTGGTGGTTGACCCAGGCAGAAGTAATTAC	2487:2516	Sγ2540:2546
	KO				
5096:5125	AATGCGCTAAACTGAGGTGATTACTCTGAG	TACAGGGG	AGACAAATCAGCTAGATTGGGCGGATCCTGG	7833:7862	Sμ4843:4849
5386:5415	CTGAGCTGGGGTGAGCTGAGCTGGGGTGAG	GG	ACCCAGACTAAATGGCTACAGAGAAGCT-A	2981:3010	
5112:5141	GTGATTACTCTGAGGTCAACAAAGCTGGGC	AGGAATCCAGTTGAG	TGGAAGAATGGGGATCCAGGCAGAGTAGCT	7989:8018	Sγ8783:8792
4817:4846	CTTAGATCCGAGGTGAGTGTGAGAGGACAG	CCCC	CCGTGTAGGGCAGCTGTAGGGAAATCAGGA	7906:7935	
4726:4755	TCTTGATCTACAACTCAATGTGGTTTAATG	GG	GGGACAGGTGGAAGTGGGAGACCCAGGCA	8203:8232	
	WT Pki				
5105:5134	AACTGAGGTGATTACTCTGAGGTAAGCAAA	TGACCCAGT	GGGTGAGCAAATACAAGGGAACTGATGGCA	2936:2965	Sμ1888:1896
5348:5377	GATCTGAAATGAGATACTCTGGAATAGCTG	GG	AACCCAGTCCAAAAACCACCAGAAGAAGCA	8450:8476	
5420:5446	AGCTGAGCTGAGCTGGGGTGATAGGTAAGA	GTGTGGGAACT	CAGTCAAAAACCACAGAAGAGCAGGAGCTA	8454:8483	Sμ4833:4840
5025:5054	GAGCCCTAGTAAGCGAGGCTCTAAAAAGCA	CAGCTG	AGTGGAGCCAAGACAGGTGGAAGTGTGGGG	2812:2841	
5349:5378	GATCTGAAATGAGATACTCTGGAGTAGCTG	CT	CCAGTTGAGGTGGAAGAATGGGGATCCAGG	7979:8008	
5043:5072	CTCTAAAAAGCACAGCTGAGCTGAGATGGG	GGC	TACAGAGAAGCTGAGGCAGGTAAGAGTGTG	8418:8447	
5304:5333	TGAGCTGGGGTGAGCTCAGCTATGCTACGC	TGTGTT	GAGTGTGGGAACCCAGTCAAAAACCACAGA	8441:8470	
5239:5267	CTTCATTAATCTAGGTTGAATAGAG-TAAA	TGGC	TACAGAGAAGCTGAGGCAGGTAAGAGTGTG	8418:8447	
5289:5318	TGTTCTGAGCTGAGATGAGCTGGGGTGAGC	AG	GACCCAGACTAAATGGCTACAGAGAAGCTG	2980:3009	
5432:5461	TAGGGTGAGCTGAGCTGGGTGAGCTGAGCT	AAAC	AAGAGTGTGGGAACCCAGTCAAAAACCACA	8439:8468	
	KO Pki				
4842:4871	GACAGGGGCTGGGGTATGGATACGCAGAAG	TGAGCTGAGTTGAGCTGGGATGAGCTGGGGTGAGCTG	AGCTCCAGGGCAGCCAGGACAGGTGGAAGT	4175:4204	Sμ5471:5507
5248:5277	TCTAGGTTGAATAGAGCTAAACTCTACTGC	TAAATGGC	TACAGAGAAGCTGAGGCAGGTAAGAGTGTG	8418:8447	Sγ2989:2996
5311:5350	TGAGCTCAGCTATGCTACGCTGTGTTGGGG	GTGATCTGA	CAGACTAAATGGCTACAGAGAAGCTGAGCA	2984:3013	Sμ5342:5350
5448:5477	GGGTGAGCTGAGCTGAGCTGGGGTAAGCTG	GG	AGGGCAGGGACCCAGACTAAATGGCTACAG	2972:3001	
5025:5054	GAGCCCTAGTAAGCGAGGCTCTAAAAAGCA	CAACTGAG	CTGATGGCAAATGGAAGGGCAGGGACCCAG	2957:2986	Sγ3176:3183
5330:5359	ACGCGTGTTGGGGTGAGCTGATCTGAAATG	TGATACTCTGCAT	ACCCAGACTAAATGGGCTACAGAGAAGCTG	2981:3009	Sμ225:231
5395:5424	GGTGAGCTGAGCTGGGCTGAGCTGGACTGA	TGGCC	TACAGAGAAGCTGAGCCAGGTTAAGAGTGG	8418:8446	
5248:5277	TCTAGGTTGAATAGAGCTAAACTCTACTGC	TAAATG	GCTACAGAGAAGCTGGGGCAGGTAAGAGTG	8416:8445	
5388:5416	GAAGATGGGGTGAGCTGAGCTGGGCTGAGC	GACCAGC	CAGAGTAGCTACAGGGAG--CAGACAGGGT	4659:4688	

Table 7. Sequence composition of >1 nucleotide CSR insertions in *Poln* mice. New mutations (different from the reference sequence) are shown in red. For some longer insertions (blue), homologies were identified in the switch region, at the positions indicated in the right column. Microhomologies at the junction site are underlined. Sequences (graphed in **Figure 21**) are from wild-type (*Poln*$^{+/+}$) and *Poln*-null (*Poln*$^{-/-}$) splenic B cells that were treated with NU7026 (Pki) or mock-treated.

Neither POLN nor POLQ alter DSB repair capacity or kinetics in the testis

Based upon knowledge that both *Poln* and *Polq* appear to be highly expressed in the testis (**Figure 20**) (Yousefzadeh et al., 2014), we hypothesized that both enzymes are involved in DNA repair in this tissue. The formation of H2AX foci and the resultant p53 signaling that occurs in mouse testis is known to occur independently of DNA-PK,

suggesting that these γH2AX foci are involved in repair pathways outside cNHEJ (Hamer et al., 2003). We hypothesize that these foci are formed during altEJ and that POLN, POLQ, or both enzymes are required for *in vivo* DSB repair in the testis of mice. Wild-type, *Poln*, and *Polq*-deficient mice were irradiated to study dose response, repair over time and check for the presence of γH2AX foci in the spermatids of mice. Spermatids are post-meiotic cells and thus are no longer undergoing programmed *Spo11*-dependent DSBs, therefore any observe γH2AX foci should reflect DNA damage inflicted by radiation or spontaneous DNA damage (Mahadevaiah et al., 2001).

γH2AX foci found in spermatids could either be the result of DSBs or replication stress (Hamer et al., 2003; Mahadevaiah et al., 2001). As expected γH2AX foci increased in a dose-dependent manner. Spermatids from the testis of *Poln* or *Polq* mice did not display an elevated level of γH2AX foci above the wild-type mouse upon x-irradiation (**Figure 22A**). γH2AX foci reached maximum levels at 30 min post-irradiation before returning to pre-irradiated levels 48 h later (**Figure 22B**). Both *Poln* and *Polq* mice behaved similar to wild-type mice. This data suggests that POLN and POLQ do not play a prominent role in the repair of x-ray-induced DSBs in mouse testis. Note that unirradiated *Polq* mice do have elevated spontaneous levels of γH2AX foci when compared to control mice (**Figure 22**). However it does not appear that POLQ-dependent altEJ plays a significant role in the repair of DSBs in the testis.

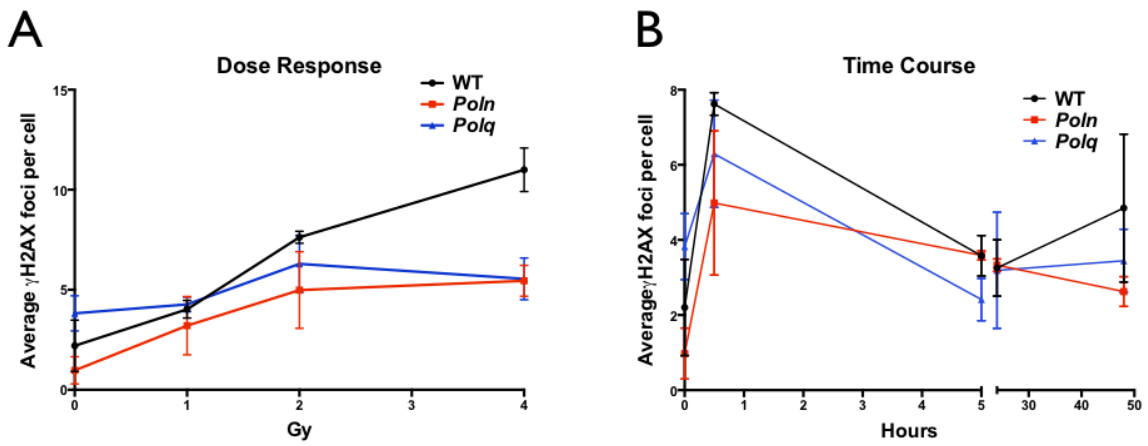


Figure 22. Loss of *Poln* or *Polq* does not dramatically alter DNA repair capacity in mouse testis. (**A**) Dose response study of wild-type (WT), *Poln*, and *Polq* mice. Mice were irradiated at the indicated Gy and were sacrificed 0.5 h later. (**B**) Time course of DNA repair. Mice were irradiated with 2 Gy and sacrificed at the indicated time points. Testis were removed post-mortem and sent for immunohistochemical processing. γH2AX foci were enumerated for 300 spermatids per mouse. 3 mice of each genotype were collected per data point, yielding a total of 900 spermatids analyzed for each experimental group. Mouse colonies were maintained by Dr. Kei-ichi Takata, Shelley Reh, and MJY. Mice were irradiated by Dr. Kei-ichi Takata and MJY and testis were harvested by Dr. Kei-ichi Takata, Shelley Reh, and MJY. Histologic preparations were made by the Science Park Histology Core and stained for γH2AX and DAPI by the Science Park Histology Core and Shelley Reh. Shelley Reh captured 63X immunofluorescence images on DMI6000B Leica fluorescence microscope. Dr. Kei-ichi Takata, Shelley Reh, Allegra Abbey, and MJY quantified foci. Data were analyzed by MJY.

Loss of Poln does not enhance DNA damage sensitivity in Polq$^{-/-}$ cells

We hypothesized that deletion of *Poln* in a *Polq*-deficient background could exacerbate the hypersensitivity to DNA damaging agents seen in *Polq*-null mouse cells (Yousefzadeh and Wood, 2013; Yousefzadeh et al., 2014). *Poln*$^{-/-}$ mice were bred to *Polq*$^{-/-}$ mice to eventually generate *Poln*$^{-/-}$ *Polq*$^{-/-}$ mice. The *Poln*$^{-/-}$ *Polq*$^{-/-}$ mice are fertile and no gross alterations to testis morphology based on histologic samples, were found (data not shown, MJY, Kei-ichi Takata, and Richard Wood). I isolated mouse embryonic fibroblast (MEF) lines from these mice and along with the appropriate control cell lines, were exposed to bleomycin and mitomycin c and analyzed for cell survival. Unlike loss of *Polq*, loss of *Poln* did not induce hypersensitivity to bleomycin (**Figure 23A**). Furthermore, deletion of both *Poln* and *Polq* did not cause a greater increase in bleomycin sensitivity when compared to the *Polq*$^{-/-}$ MEFs (**Figure 23A**). These data suggest that POLN and POLQ are not epistatic to one another and that POLN does not participate in the same repair pathway as POLQ. *Poln* and *Polq* single or double knockout MEFs did not exhibit hypersensitivity to the ICL-forming agent mitomycin c (MMC) (**Figure 23B**). Lack of MMC sensitivity suggests that neither POLN nor POLQ function in the HR pathway to repair ICLs.

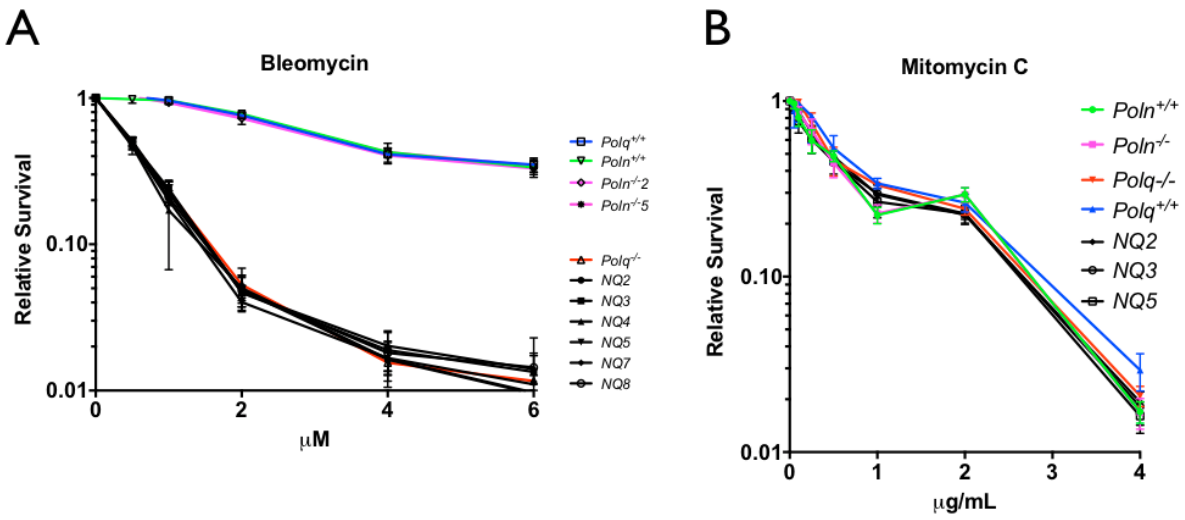


Figure 23. Loss of *Poln* does not enhance DNA damage hypersensitivity found in *Polq* MEFs. *Poln* and *Polq*-proficient and -deficient MEFs, as well as the newly created *Poln Polq* double KO MEFs (NQ) were plated in 96 well plates and exposed to either (**A**) bleomycin or (**B**) mitomycin c and cellular ATP levels were measured as a surrogate for cell survival. Dr. Kei-ichi Takata and Shelley Reh derived and established the *Poln* MEFs. The *Polq* MEFs and *Poln Polq* MEFs were derived and established by MJY. MJY performed damage sensitivity assays.

As discussed in Chapter 4, multiple studies have shown that cells lacking POLQ are hypersensitive to ionizing radiation and DSB-inducing compounds. POLQ confers resistance to DNA damage through its polymerase activities and participates in an altEJ pathway (**Figure 7**, **Figure 11-12**) (Yousefzadeh et al., 2014). Overexpression of *POLQ* in the tumors of breast cancer patients also confers poorer patient outcomes (Begg, 2010; Higgins et al., 2010a; Lemée et al., 2010). Coupling these facts together, I hypothesized that the aberrant levels of POLQ observed in these breast cancer patients contributes to their poor outcomes by making them resistant to DNA damage, possibly through elevated altEJ activity. To test the hypothesis that aberrant levels of POLQ induces DNA damage resistance, 293T cells were transfected with POLQ expression constructs to express either wild-type POLQ, polymerase-dead POLQ, or an empty vector control and assayed for bleomycin sensitivity (**Figure 24**). While empty vector control or polymerase-dead POLQ failed to confer additional bleomycin resistance, ectopic expression of wild-type POLQ conferred increased resistance to bleomycin. This finding suggests that ectopic levels of POLQ are mediating DNA damage resistance in cancer cells through use of its polymerase activity and that POLQ may be the rate limiting factor in POLQ-dependent altEJ.

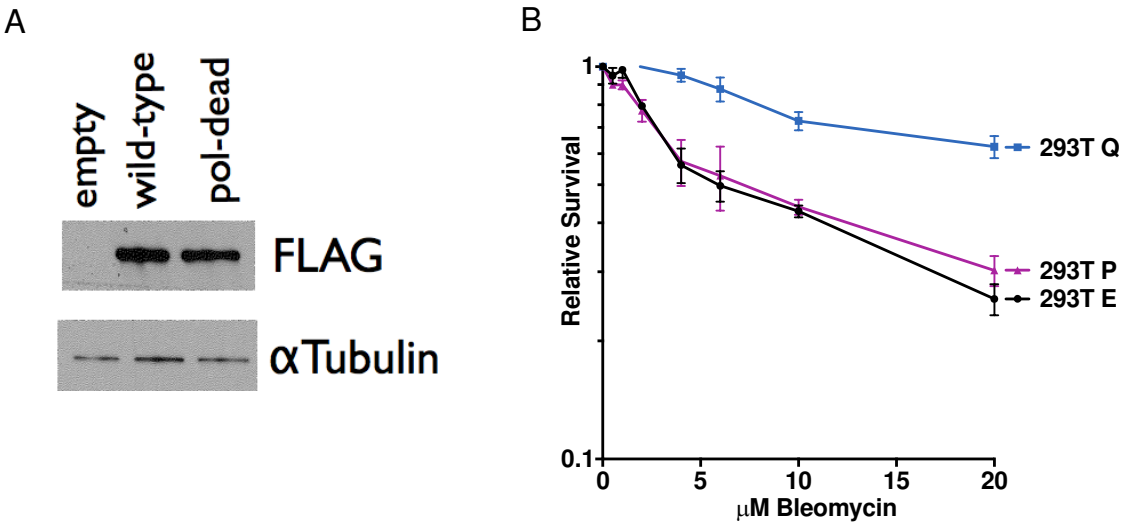


Figure 24. Ectopic expression of POLQ in cancer cells induces DNA damage resistance. (**A**) 293T cells were transiently transfected with pCDH containing no insert (E), wild-type POLQ (Q), or polymerase-dead POLQ (P) using jetPRIME (Polyplus) according to manufacturer's specifications. Cells were then plated in triplicate in 96 well plates and exposed to the indicated concentrations of bleomycin for 72 h. Cellular ATP levels were measured for three independent experiments as a surrogate for survival. (**B**) Representative immunoblot of ectopic POLQ levels at the time of damage. αTubulin serves as a loading control. Cell culture, transfections, immunoblotting, and damage sensitivity assays were performed by MJY.

5.4 Conclusion and future directions

The data presented here show that consistent with previous reports, *Poln* is highly expressed in the testis (**Figure 20**) (Marini et al., 2003). This suggests that any *in vivo* function for POLN may be detectable only be in the testis. Loss of *Poln* did not impair class switch recombination or alter how the breaks are repaired in B cells (**Figure 21**). While previous observations demonstrate that longer insertions in the CSR junctions of B cells are POLQ-dependent (**Figure 11**) (Yousefzadeh et al., 2014), POLN appears to be dispensable for the formation of these long insertions in this cell type (**Figure 21**, **Table 7**). Although the polymerase domain of POLN shows homology to those of POLQ and *Drosophila* Mus308, it does not appear to operate in the altEJ pathway they do (**Figure 21**, **Table 7**) (Chan et al., 2010; Yousefzadeh et al., 2014; Yu and McVey, 2010). Furthermore, mice lacking *Poln* do not show an increase in DNA damage signaling or a persistence of markers of DSBs after irradiation (**Figure 22**). Mice lacking *Polq* do show an increase in DNA damage markers, but only spontaneously and not upon exposure to ionizing radiation (**Figure 22**). This finding is consistent with previous work showing inherent genome instability as measured by micronuclei formation that spontaneously occur both in the cultured cells and the peripheral blood of mice lacking *Polq* (Shima et al., 2004; Yousefzadeh et al., 2014). The lack of enhanced or persistent DNA damage signaling in spermatids in the absence of POLQ, could be explained by the fact that the POLQ-dependent SD-EJ pathways does not participate in the repair of DSBs that occur in murine testis.

Prior reports provided evidence of ICL hypersensitivity upon suppression of POLN by siRNA in human cancer cell lines (Moldovan et al., 2010; Zietlow et al., 2009). In contrast with these published reports, we show that MEFs lacking *Poln* do not display hypersensitivity to MMC (**Figure 23B**). This could be a cell-specific effect whereby MEFs may not show hypersensitivity if they do not form full-length POLN (the majority of transcripts could be non-functional splicing variants) but POLN may still play a role in the

repair of ICL-mediated DNA damage. Loss of *Poln* did not enhance the DNA damage sensitivity phenotype seen in *Polq*-null cells and loss of *Poln*, *Polq*, nor did both genes did not affect MMC sensitivity (**Figure 23**). These results provide evidence that POLN and POLQ are not epistatic with each other and that POLN does not function in the same altEJ pathway as POLQ does. Based upon tissue expression data (**Figure 20**) and a suggested function for POLN in homology-directed repair (Kohzaki et al., 2010), a role for POLN in meiotic recombination may exist, but any evidence is lacking.

We have yet to find an unequivocal phenotype for mice lacking *Poln*, this may be because *Poln* mice have an extremely subtle phenotype, or the redundancy of other specialized DNA polymerases (or other DNA repair pathways) in mammals masks the effects of *Poln* deficiency. The original intent of crossing the *Poln* mice to the *Polq* mice was that loss of *Poln* would further stress the *Polq* mice and allow insight into the *in vivo* function of POLN. Unfortunately, this exacerbation of the *Polq* phenotype did not occur and leaves us with no phenotype for the *Poln* mice. Though they are related DNA polymerases, POLN and POLQ do not appear to have similar functions in the cell. Crossing the *Poln* mice to other mutants like *Rev1* mice may give insight as to POLN's function in the cell.

POLN was previously reported to interact with components of the Fanconi anemia (FA) pathway (Moldovan et al., 2010) and Drs. Kei-ichi Takata and Junya Tomida found interaction between POLN and FANCJ (unpublished data, Drs. Kei-ichi Takata, Junya Tomida, and Richard Wood). While we have been unable to detect hypersensitivity of POLN to ICL agents, the interaction with FANCJ may still be explained in a plausible fashion (**Figure 23B**). It is known that FANCJ is quite adept at unwinding G-quadraplexes (G4) and this activity occurs independently of the rest of the FA pathway (Brosh and Cantor, 2014). FANCJ could be opening up these G4s for specialized DNA polymerases (like POLN) to synthesize through these structures. The lack of a significant detectable phenotype in *Poln* knockouts could be explained through functional redundancy of the polymerases.

For example, REV1 is the primary DNA polymerase involved in replicating past difficult G4-mediated replication blocks in chicken DT40 cells (Sarkies et al., 2012); therefore, deletion of *Rev1* would lead to the cell using other specialized DNA polymerases like POLN in order to replicate G4 sites in DNA. If so, then a severe replicative stress phenotype might be seen in mice that lack both *Rev1* and *Poln*.

While a role for POLN in the repair of ionizing radiation-induced DNA damage has yet to be substantiated, one clearly exists for POLQ. As discussed in Chapters 2 and 4, POLQ mediates the repair of DSBs through participation in an altEJ pathway and overexpression of POLQ in breast cancer confers poorer outcomes (Lemée et al., 2010; Yousefzadeh and Wood, 2013; Yousefzadeh et al., 2014). Preliminary evidence presented in this chapter demonstrates that aberrantly high levels of POLQ confer DNA damage resistance to cancer cells (**Figure 24**). This could mimic what is occurring in these breast cancer patients with elevated *POLQ* levels, suggesting that chemoresistance may be the mechanism that is conferring their poorer outcomes, but many other genes are also co-upregulated in cancer. Further analysis is needed to understand the detailed mechanism of how POLQ-dependent end joining promotes DSB repair and why elevated *POLQ* expression has a negative impact on patient outcomes in breast cancer.

Chapter 6. Generation and validation of human POLQ antibodies.

Note: This chapter discusses the generation of human POLQ antibodies that was discussed in Chapter 2 and used in subsequent chapters, in addition to other antibodies that were used in the laboratory.

6.1 Introduction

A valuable tool for studying the function of POLQ in cells is an antibody that recognizes endogenous POLQ with great specificity. Until recently, an antibody that recognizes endogenous POLQ in immunoblots or for immunoprecipitation studies, was unavailable (Fernandez-Vidal et al., 2014). The Wood laboratory previously produced a pair of rabbit polyclonal antibodies (NT and CEN) that were raised against fragments of recombinant human POLQ (Seki et al., 2003), but the supply of these two antibodies was exhausted. Our lab endeavored to generate a new series of antibodies against POLQ and set upon making monoclonal antibodies in order to create a POLQ antibody that had great specificity.

6. 2 Material and Methods

Please note that this section is written as a protocol for persons that may be trying to repeat these experimental procedures.

Note on centrifugation:

Cells were centrifuged in Thermo Sorvall Legend RT+ in Sorvall Heraeus swinging bucket rotor at 1000 or 3000 rpm. At the bench top, lysates were centrifuged in an Eppendorf 5424 R centrifuge at 15000 rpm / 21,000xg.

Note on lysates:

All lysates were mixed 1:1 with 2X SDS loading buffer and heat denatured at 95℃ for 5 min.

Western Blot Protocol

Use 4-20% TGX gels from Bio-Rad

-Run gel at 150 V for 1-1.5 h.

-Transfer to PVDF Immobilon at 20 V overnight or 30 V for 3 h in transfer buffer (1X Tris-Glycine + 10% methanol) at 4° C.

-Block 1 h in 10% milk TBST.

-Incubate in primary antibody overnight (in TBST, no milk) at the indicated concentration at 4° C. Antibodies are also stored in 1X TBST with 1:400 dilution of 5% sodium azide.

-Wash three times in 1X TBST for 5 min.

-Incubate in secondary anti-mouse (A0168) or anti-rabbit (A0545) antibody (Sigma-Aldrich) (1:10,000 dilution in 5% milk TBST) for 1.5-2 h at room temperature.

-Wash three times in 1X TBST for 5 min.

-Visualize with ECL.

Nuclear Extract Preparation

Cells are trypsinized and neutralized with media.

-Pellet cells in Sorvall, 1000 rpm, 5 min, 4°C.

-Wash pellet with PBS. Count cell number, need at least 2 x10^8 cells for extracts, but can make larger extract preps.

-Pellet cells in Sorvall, 1000 rpm, 5 min, 4°C.

-Wash pellet with Hypotonic Buffer.

-Pack cells in Sorvall, 3000 rpm, 5 min, 4°C.

-Add 2 packed cell volumes of Hypotonic Buffer and resuspend cells.

-Homogenize 10X **slowly** in 7 mL Wheaton Dounce Homogenizer (Tight). Check lysis by trypan blue staining.

-Centrifuge in Sorvall 13000 rpm, 15 min, 4°C.

-Collect supernatant as cytosolic fraction. _____ mg/mL

-Wash pellet in 5-10 volumes of Hypotonic Buffer.

-Pellet cells in Sorvall 13000rpm, 5 min, 4°C.

-Resuspend in 2 volumes of Nuclear Extract Buffer.

-Homogenize 10X slowly in 7 mL Wheaton Dounce Homogenizer.

-Incubate on rotation for 30".

-Spin 13000 rpm, 15 min, 4°C.

-Collect supernatant as NE1. _____ mg/mL

-Store extract in minus 80 freezer.

Hypotonic Buffer	Nuclear Extract Buffer
10 mM HEPES, pH 7.9	10 mM HEPES, pH 7.9
10 mM KCl	0.42 M NaCl
1.5 mM $MgCl_2$	1.5 mM $MgCl_2$
0.5 mM PMSF	0.5 mM PMSF
1 mM DTT	1 mM DTT
	10% Glycerol
Roche Protease Inhibitor Cocktail Tablet	Roche Protease Inhibitor Cocktail Tablet

Notes:

Roche Protease Inhibitor Cocktail Tablet (Catalog # 11 836 170 001)

Adapted from (Dignam et al., 1983; Lee et al., 1988).

Whole Cell Extract Preparation

-Resuspend 5 x10^6 cells in 40 µL of Lysis buffers and incubate on ice in the cold room for 30 minutes.

-Centrifuge 13,000 rpm, 15 min, 4℃

-Collect supernatant as whole cell extract.

Whole Cell Extract Buffer

50 mM Tris-HCl pH 7.5

0.25 M NaCl

1 mM EDTA

0.5 mM PMSF

0.1% Triton X-100

Roche Protease Inhibitor Cocktail Tablet

Protocol adapted from (Welsh et al., 2004).

Transfection of 293T cells and Crude SDS Lysate Preparation

-1.5 x10^5 293T cells were plated into each well of a 6-well dish.

-The following day, cells were transfected with 2 µg of either pCDH/empty vector control or pCDH/POLQ.

-48 h post-transfection cells from one well of a 6 well plate were resuspended in 200 µL of 2X SDS Loading Buffer.

-Heat denature at 95℃ for 5 min.

Immunoprecipitation Protocol

-Plate 1.5 x10^6 293T cells into four 10 cm dishes with 10 mL of complete media (DMEM).

-The following day, cells were transfected with 10 µg of either pCDH/empty vector control or pCDH/POLQ.

-48 h post-transfection, scrape cells in 1X PBS. Centrifuge 3,000 rpm, 5 min, 4℃.

-Resuspend cells in 200 µL per dish in IP Lysis Buffer. Vortex for 30 seconds and agitate in the cold room for 15 minutes. Centrifuge 13,000 rpm, 15 min, 4℃.

-Collect supernatant. Determine concentration. _____ mg/mL

-Collect 25 µL and add 1:1 volume of 2X SDS Lysis Buffer. Heat denature at 95℃ for 5 m.

-Wash 50 µL (per sample) FLAG-Dynabeads three times in IP Buffer.

-Use 5 mg of protein in 2 mL total volume for IP. Incubate overnight on rotation in cold room.

-Wash 3X in IP Wash Buffer.

-Add 50 µL of 2X SDS Lysis Buffer to beads and heat denature at 95℃ for 5 m.

-Transfer liquid to new tube and freeze.

-For immunoblot: Load 10 µL of input (30 g total protein) and 25 µL IP product.

IP Lysis Buffer	IP Buffer	IP Wash Buffer
50 mM Tris-HCl pH 7.5	50 mM Tris-HCl pH 7.5	50 mM Tris-HCl pH 7.5
0.15 M NaCl	0.15 M NaCl	0.15 M NaCl
1 mM EDTA	1 mM EDTA	1 mM EDTA
0.5 mM PMSF	0.5 mM PMSF	0.5 mM PMSF
1% Triton X-100	0.2% NP-40	0.2% NP-40

Please add a Roche Protease Inhibitor Cocktail Tablet to each buffer before use.

Adapted from protocol of Arijit Dutta of Sankar Mitra's Lab (Houston Methodist Research Institute).

Immunofluorescence

-1.5 x10^4 293T cells were plated into 4-well chamber slides (LabTek).

-The following day, cells were fixed in 2% p-formaldehyde for 10 min at room temperature.

-Rinse slides with 1X PBS.

-Permeabilize cells in 1X PBS + 0.2% Triton X-100 for 2 min at room temperature.

-Block for 30 m in 1X PBS + 0.2% Triton X-100 + 5% BSA at room temperature.

-Dilute primary antibody to 1:100 in 1X PBS + 0.2% Triton X-100 + 5% BSA and incubate overnight at 4℃.

- Wash slides in 1X PBS + 0.2% Triton X-100 for 5 min at room temperature.

- Dilute secondary antibody (Molecular Probes Anti-mouse AlexaFluor-488) to 1:1,000 in 1X PBS + 0.2% Triton X-100 + 5% BSA and incubate 1 h at room temperature.

- Wash slides in 1X PBS + 0.2% Triton X-100 for 5 min at room temperature.

-Mount with Vectashield + DAPI (Vector Labs) [Note dilute Vectashield + DAPI 1:5 in Vectashield without DAPI. Or better yet, just DAPI stain and wash to clear up heavy DAPI background.].

-Paint edges of slides with nail polish.

Adapted from (Bhagwat et al., 2009).

<u>2X SDS Loading Buffer</u>

20% Glycerol

4% SDS

0.2% Bromophenol Blue

12% 2-Mercaptoethanol

100 mM Tris-HCl pH 6.8

6.3 Results

Matthew Hogg of Sylvie Doublié's laboratory (Univ. of Vermont) graciously provided us with a fragment of recombinant POLQ (**Figure 25A**) that possessed polymerase activity (Hogg et al., 2011). This recombinant fragment was expressed in *E. coli* and was used to generate antibodies against POLQ. The MDACC antibody core and Epitomics respectively generated the mouse and rabbit monoclonal antibodies, respectively. A SUMO3 fusion tag protein made from the same vector (Life Sensors) as the recombinant fragment was sent to both sites as a control protein for ELISA screening. Over 80 mouse hybridoma clones that were positive by ELISA were provided to me for further screening (**Figure 25B**). I immunoblotted the recombinant protein fragment used to immunize the mice with the supernatants from the hybridoma clones (**Figure 25B**) and narrowed the selection down to 13 different mouse monoclonal antibodies that underwent further purification and analysis (**Table 8**). The rabbit monoclonal antibodies were screened in a similar fashion (**Figure 25C**). The purified mouse monoclonal antibodies were then screened against whole cell and nuclear extracts from various cell lines (**Figure 25D-F**) as well as recombinant full-length POLQ (**Figure 25G**).

Each monoclonal antibody recognizes a band larger than 250 kDa that might represent full-length POLQ. Attempts to determine whether any of these bands are endogenous POLQ by siRNA suppression have yielded variable results (**Figure 25H**). Each of the monoclonal antibodies was epitope mapped by Shelley Reh using peptide arrays from JPT (**Figure 26**, **Table 8**). The POLQ20 antibody was predominantly used for immunoblotting in the laboratory because it maps near motif 1 and insert 1 of the polymerase domain of POLQ (**Figure 26**, **Table 8**), and provides high sensitivity and low background on immunoblots. It may be useful for immunoprecipitation experiments (**Figure 25F-I**). The POLQ20 antibody is able to recognize recombinant human POLQ, high molecular weight bands (>250 kDa) in whole cell extracts that could be consistent with

endogenous POLQ. POLQ20 recognizes no bands that correspond to full-length POLQ when immunoblotted against extracts made from MEFs (**Figure 27**). All mouse monoclonal hybridoma cell lines are stored at the MDACC monoclonal antibody core facility in Houston and purified POLQ monoclonal antibodies are stored in my -80° C storage. A monoclonal antibody against POLQ (pol θ 1B1) provided by Jean-Sebastién Hoffman was screened with various extracts and recognizes a band very similar to the POLQ20 (**Figure 28**) (Fernandez-Vidal et al., 2014). This antibody is able to recognize recombinant POLQ but recognizes higher molecular weight bands sporadically upon very high levels of exposure, making it a less than ideal antibody for use.

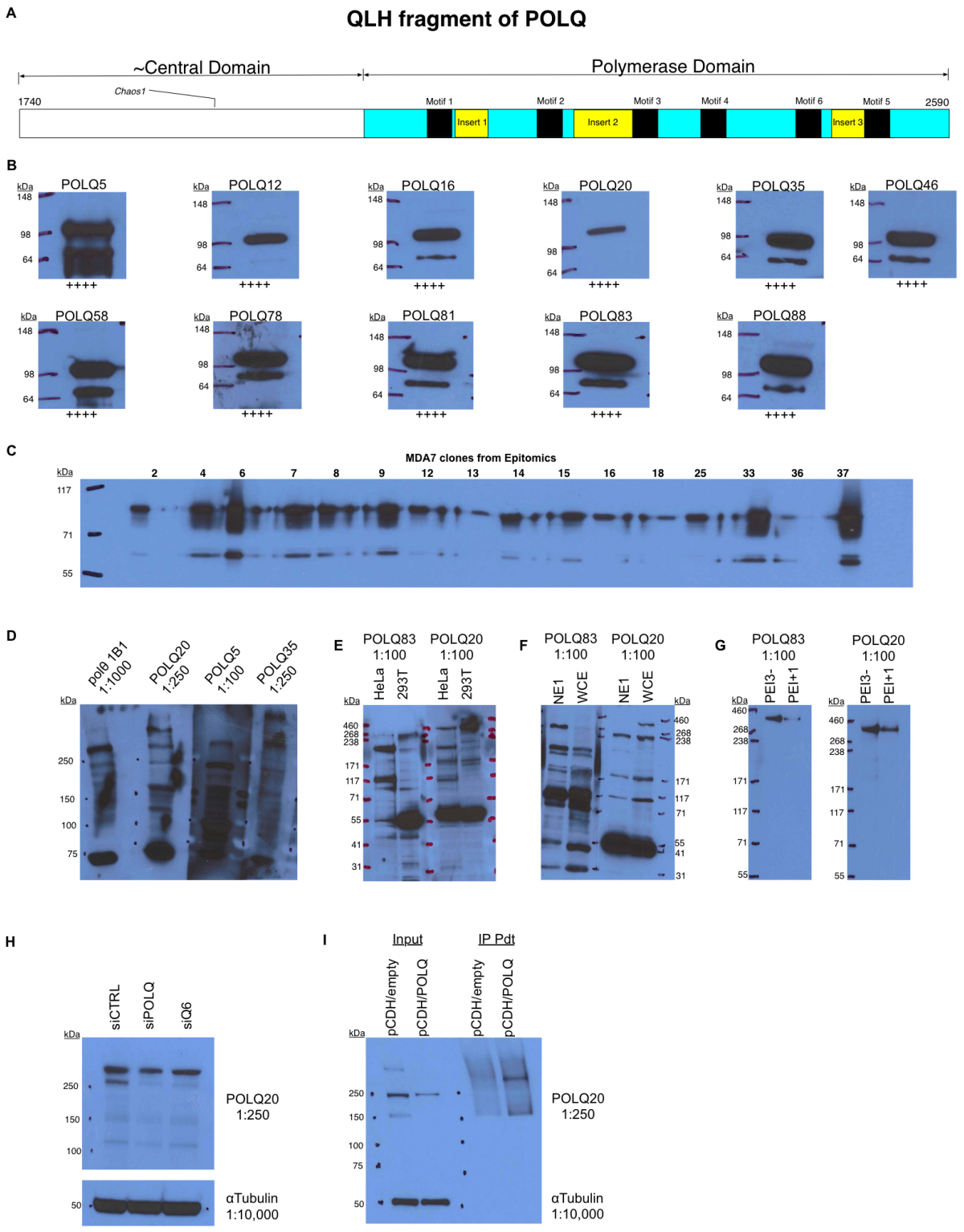


Figure 25. Screening of human POLQ antibodies. (**A**) Schematic of the QLH fragment of POLQ that was used to immunize rabbits and mice to generate monoclonal antibodies. (**B**)

Primary screening of raw mouse monoclonal supernatants by immunoblotting against 150 ng of purified QLH fragment. + symbols below immunoblots represents the initial ELISA analysis of monoclonal antibodies (++++ being the highest score and + being the lowest score. (**C**) Rabbit monoclonal supernatants from Epitomics were similarly screened. (**D**) Purified monoclonals, including one from Jean-Sebastién Hoffman are immunoblotted against 75 µg of RKO whole cell extracts (WCE). POLQ83 and POLQ20 mouse monoclonal antibodies were screened against (**E**) 100 µg WCE from HeLa and 293T cells, (**F**) 35 µg nuclear extract 1 and 100 µg WCE from 293T cells, and (**G**) 10 µL of recombinant POLQ preps purified from baculovirus-infected insect cells. (**H**) RKO cells were transfected with non-targeting siRNA (siCTRL), a pool of 4 siRNAs (siPOLQ), or a single siRNA (siQ6) against POLQ. WCE were made 48 h post transfection and 100 µg were immunoblotted with the POLQ20 mouse monoclonal antibody. The band below just above the 250 kDa marker is putatively thought to be POLQ. (**I**) Immunoblot of FLAG-IP experiment with 293T cells transiently transfected with either pCDH/POLQ or empty vector control. Arrow denotes full-length POLQ. All gels are 4-20% polyacrylamide gels (**E** is a 5% polyacrylamide gel).

Antibodies	**Epitope**	**Amino Acid #**
MOUSE MONOCLONAL ANTIBODIES[1]	*(MDACC ANTIBODY CORE)*	
POLQ20*	(REMKN)	2147:2151
POLQ5	(QPGESYLFGS)	1767:1776
POLQ12	(QPGESYLFGS)	1767:1776
POLQ16**	(PRDFEIKMP)	2246:2254
	(LNTGADVFRS)	2345:2354
POLQ35	(VNPWK)	1757:1761
POLQ46	(TSASK)	1742:1746
	(QPGESYLFGS)	1767:1776
POLQ58	(VNPWK)	1757:1761
POLQ78	(VNPWK)	1757:1761
POLQ81	(VNPWK)	1757:1761
POLQ83**	(PRDFEIKMP)	2246:2254
	(LNTGADVFRS)	2345:2354
POLQ88	(VNPWK)	1757:1761

Hoffman antibody from 2014 Nature Communications

	Antigen	
pol θ 1B1	(CSIFRARKRASLDINKEKPG)	983:1002

RABBIT MONOCLONAL ANTIBODIES[2] *(EPITOMICS)*

MDA7-8-4	(N/A)	
MDA7-8-8	(N/A)	
MDA7-8-12	(N/A)	
MDA7-14-6	(N/A)	
MDA7-14-7	(N/A)	
MDA7-14-8	(N/A)	
MDA7-15-2	(N/A)	
MDA7-15-7	(N/A)	
MDA7-15-8	(IGFSTAECESQKHIM)	2092:2106
MDA7-16-2	(N/A)	
MDA7-16-6	(N/A)	
MDA7-16-8	(N/A)	
MDA7-36-2	(N/A)	
MDA7-36-7	(N/A)	
MDA7-36-8	(N/A)	
MDA7-37-7	(N/A)	
MDA7-37-8	(N/A)	
MDA7-37-10	(N/A)	

RABBIT POLYCLONAL ANTIBODIES[3] *(CELL SIGNALING TECHNOLOGIES)*

CST PP3***
CST PP16****
CST PP27***
CST PP28***
CST PP36*****

RABBIT POLYCLONAL ANTIBODIES[4] *(BETHYL LABORATORIES)*

BL18474
BL18475
BL18476
BL18477******

[1]I typically use the mouse monoclonal antibodies at a dilution between 1:250-1:500 for western blotting and the pol θ 1B1 mouse monoclonal antibody and the rabbit polyclonal antibodies are recommended for use at 1:10,000 and 1:1000 dilutions. Epitopes to human POLQ are listed above, amino acid in red denote deviations from homology between human and mouse POLQ.
*Best antibody for western blotting, it is the most sensitive of the panel and has the least background.

**Both POLQ 16 and POLQ83 epitope map to where the secondary antibody control binds on the QLH fragment but with much greater affinity.

[2]Rabbit monoclonal hybridoma lines that were raised against the QL fragment by Epitomics are currently stored in liquid nitrogen storage on site and with Cryogene. All of the hybridomas recognize a high molecular weight band in 293T cells that could be consistent with full length POLQ. Of these antibodies MDA7-15-8 was selected and sent to the MDACC antibody core for ascites production and purification and epitope mapping. The rest of the rabbit monoclonal antibodies still need to be epitope mapped for future use.

Epitomics clones are described by Project #-Original hybridoma clone #-Subclone #. Therefore for MDA7-15-8, it is the 8th subclone of original clone 15 for the MDA7 project. These hybridomas were generated from a single rabbit.

[3]Rabbit polyclonal antibodies were generated by Cell Signaling Technologies using multiple unknown peptide sequences of POLQ. Purportedly PP36's epitope is conserved between human and mouse. PP36 recognizes recombinant POLQ but with a lower affinity when compared to the POLQ20 mouse monoclonal antibody. All of the antibodies recognize recombinant POLQ seen in lysates from 293T cells transfected with either empty pCDH vector or pCDH/POLQ by MJY. As of July 2014 the rabbits producing the PP16 and PP36 were moved into the monoclonal antibody production pipeline.
***PP3, PP27, and PP28 appear to be unsuitable as they fail to recognize bands above ~250 kDa in 293T and HeLa extracts.
**** PP16 antibody has a high background.
*****PP36 recognized a high molecular weight band that may be endogenous POLQ in 293T cells but not in HeLa cells.

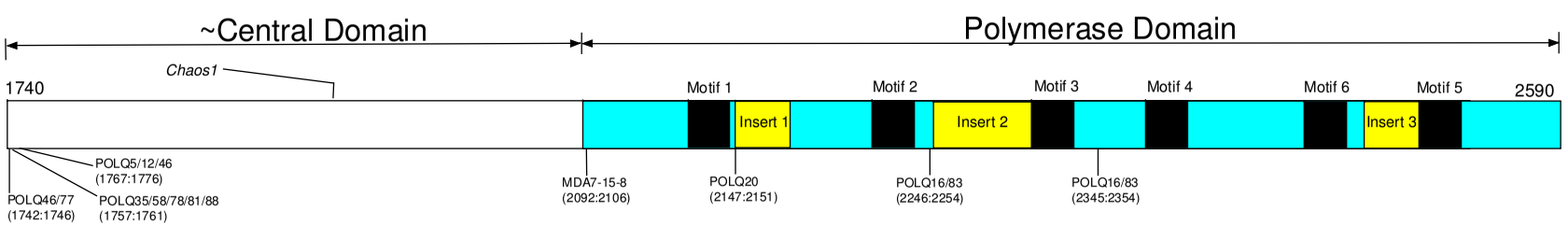
[4]Rabbit polyclonal antibodies were generated by Bethyl Laboratories using multiple unknown peptide sequences of POLQ. Purified antibody was sent to us and used at a 1:5,000 dilution for immunoblotting.
******BL18477 did not recognize recombinant POLQ in lysates from 293T cells transfected with pCDH/POLQ. I would not recommend using this antibody.

Table 8: Human POLQ antibodies. Epitope sequences and amino acid numbers are shown for each antibody. Amino acid residues in red are those that differ between human and mouse.

A

B

C



Hoffman polθ-1
(983:1002)
immunoblotting

Hoffman polθ-2
(1714:1733)
immunoprecipitation

Figure 26. Epitope mapping of human POLQ antibodies. JPT epitope mapping microarray containing peptides that correspond to the antigen sequence is diagramed (**A**) where 11 amino acid peptides are covalently attached to slide. (**B**) These peptides overlap by 7 amino acids (shown in red) with the preceding peptide in order to specifically determine the epitope site for each of the antibodies that were mapped. (**C**) Schematic of the QLH fragment of POLQ showing the location of where each monoclonal antibody maps. Amino acid numbers of epitope sites are shown in parentheses with the corresponding antibody. Epitope mapping was performed and analyzed by Shelley Reh and MJY.

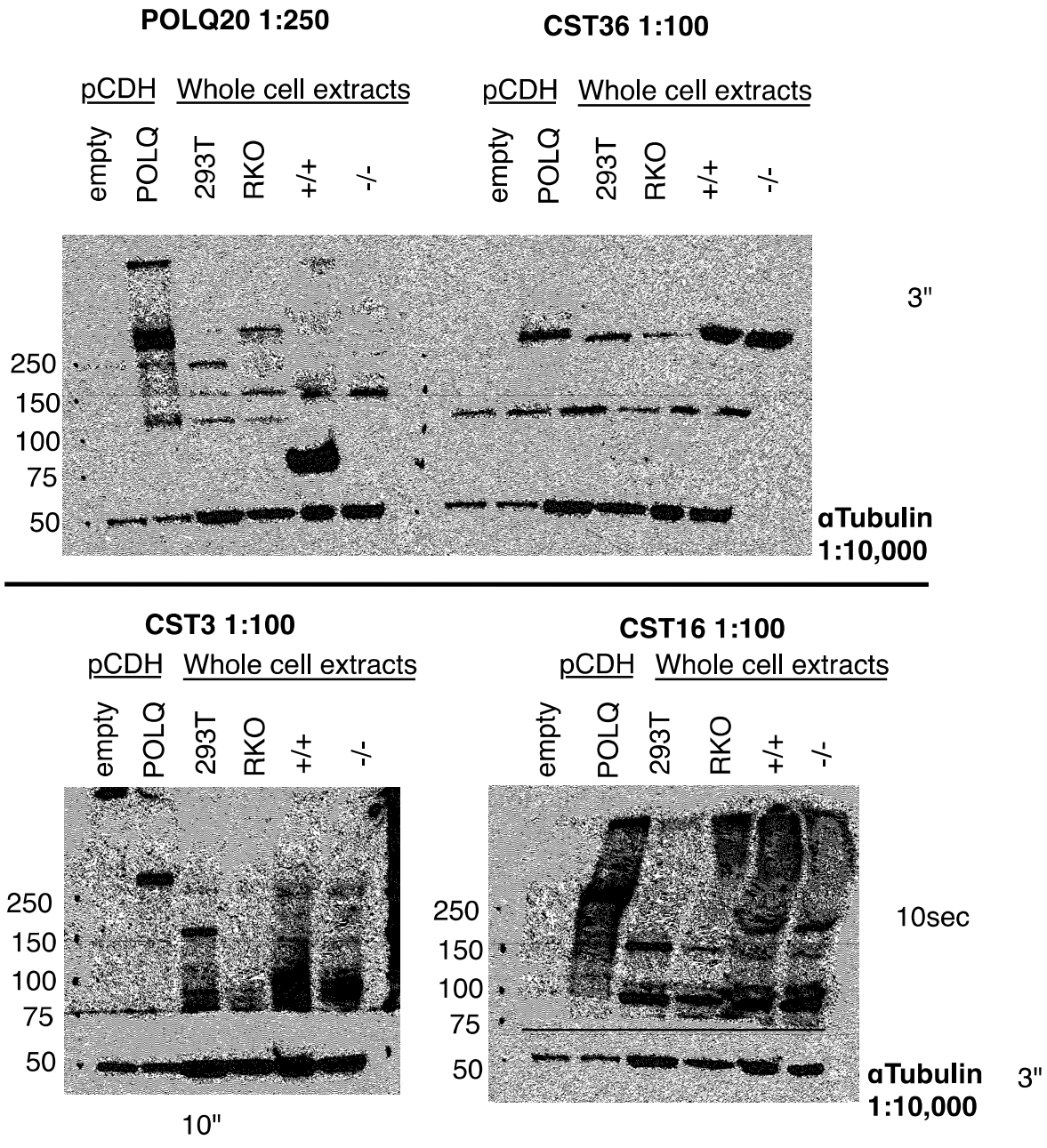


Figure 27. Screening of human POLQ rabbit polyclonal antibodies from Cell Signaling Technologies. POLQ20 mouse monoclonal antibody and three rabbit polyclonal antibodies raised against POLQ peptides (CST3, CST16, CST36) were immunoblotted with: 10 µL of crude SDS lysates from 293T cells transfected with either pCDH/empty vector control or

pCDH/POLQ; 100 µg WCE from 293T, RKO, *Polq*-proficient (+/+) and deficient (-/-) mouse embryonic fibroblasts.

Recently we have collaborated with Cell Signaling Technologies (CST) on human POLQ antibodies (**Table 8**). We received multiple rabbit polyclonal antibodies raised against different POLQ peptides for validation (CST PP 3, 16, 36). Each of the three antibodies recognized recombinant POLQ, as did the control mouse monoclonal POLQ20 antibody (**Figure 27**). The antibodies also recognized bands above 250 kDa in whole cell extracts but were unable to recognize a band specific to endogenous mouse POLQ. As of September 2014, CST was utilizing our data to move the antibodies into their monoclonal antibody production pipeline.

Human POLQ rabbit polyclonal antibodies from Bethyl Laboratories were also screened in the laboratory (**Table 8**). Three of the four antibodies provided by Bethyl Laboratories recognized recombinant POLQ (**Figure 28**), as did the mouse monoclonal antibody controls (POLQ20, pol θ 1B1). The Bethyl antibodies yield a high amount of background using the specified conditions and would need further purification (affinity purification with recombinant POLQ) or dilution before further use.

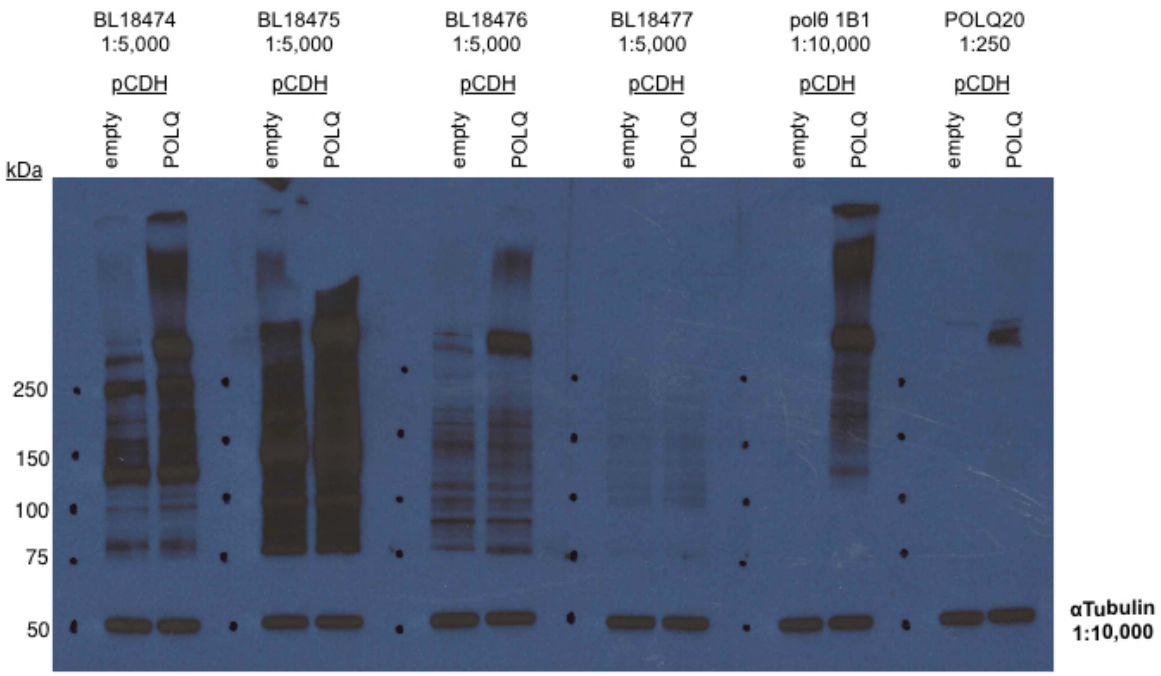


Figure 28. Screening of human POLQ rabbit polyclonal antibodies from Bethyl Laboratories. POLQ20 and pol θ 1B1 mouse monoclonal antibodies and four rabbit polyclonal antibodies (BL18474-7) raised against POLQ peptides by Bethyl Laboratories were immunoblotted with 3 µL of crude SDS lysates from 293T cells transfected with either pCDH/empty vector control or pCDH/POLQ.

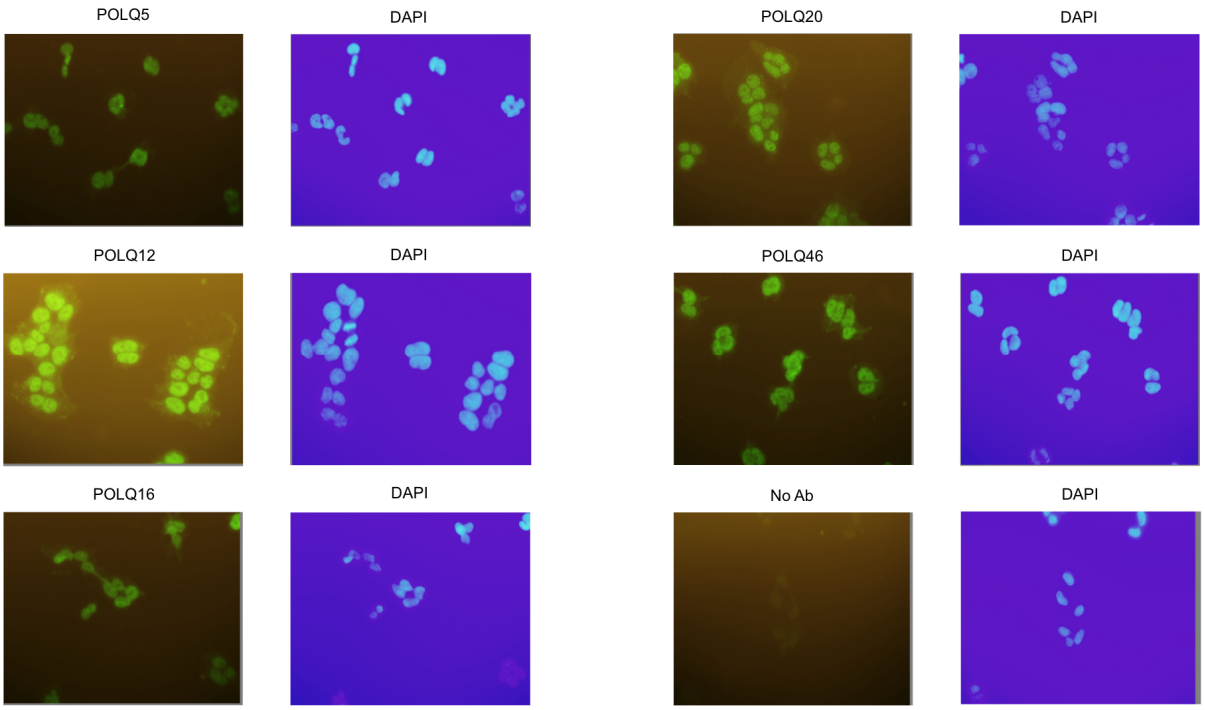


Figure 29. Cell staining and immunofluorescence with human POLQ antibodies. 293T cells were fixed and stained with the indicated mouse monoclonal antibodies and AlexaFluor 488 conjugated secondary before stained with DAPI and imaged.

6.4 Conclusion and future directions

Moving forward, there are many applications for which these antibodies against POLQ could be of great use. These antibodies would be useful for immunofluorescence (**Figure 29**), immunoprecipitation (**Figure 25I**), and potentially immunohistochemical staining of human tissues. However, further validation of these antibodies is needed and proper POLQ human knockout cell line would be vital to confirm specificity of these antibodies. Dr. Kei-ichi Takata and Mrs. Shelley Reh (Wood laboratory) are currently generating POLQ human knockout cell lines. These cell lines could be the key reagent to screening the POLQ antibodies to find an antibody clone that recognizes POLQ with great sensitivity and specificity. Then a validated antibody could be used to further determine subcellular localization of POLQ, its interaction partners, and potentially any relevant post-translational modifications that modulate its activity.

As it stands, the POLQ20 mouse monoclonal antibody appears to be the most suitable for detecting human POLQ. This is based upon the high degree of sensitivity and modest background levels observed when immunoblotting with this antibody as well as its ability to recognize recombinant human POLQ and high molecular weight bands in an extract (**Figure 25, Figure 27**). Only POLQ20 and POLQ83 mouse monoclonal antibodies have been used in the immunoblotting of extracts made from MEFs. While both antibodies recognize high molecular weight bands, they do not appear to be specific to POLQ. The possibility that one of other mouse monoclonal antibodies against POLQ that were generated by the MD Anderson Monoclonal Antibody Core recognizes endogenous mouse POLQ is possible as they have not been immunoblotted with extracts from the MEFs. My frank suggestion is avoid using the pol θ 1B1 mouse monoclonal antibody. It does not reliably recognize high molecular weight bands in whole cell extracts of *Polq* MEFs or HCT116 cells. Until a POLQ antibody has been properly vetted we must maintain a healthy dose of skepticism with regard to immunoblotting results with antibodies against POLQ

Chapter 7. Future directions.

7.1 POLQ maintains genomic integrity through its altEJ activity

Genetic screens have shown that POLQ and its homologs in *C. elegans* and *Drosophila* are key components of an alternative end joining pathway that facilitates the repair of DNA double strand breaks (DSBs) (**Figure 5**, **Figure 12**) (Koole et al., 2014; Yousefzadeh et al., 2014; Yu and McVey, 2010). This repair pathway, termed synthesis-dependent end joining (SD-EJ) or synthesis-dependent microhomology-mediated end joining (SD-MMEJ), relies upon POLQ-dependent DNA polymerization to extend enzymatically-processed 3' DNA ends to create a novel microhomology. This homologous region can then anneal to the other 3' end from the DSB (Yousefzadeh et al., 2014). This complex nucleotide insertion relies upon the presence of a free 3' end (whether it be from the other end of the break or from another free DNA end) to provide a suitable template for creating the "insertion" (Chan et al., 2010; Yousefzadeh et al., 2014; Yu and McVey, 2010). This alternative end joining pathway appears to be suppress to the formation of large deletions and chromosomal translocations (Roerink et al., 2014; Yousefzadeh et al., 2014). Consistent with this idea, we observed a four-fold increase in the frequency of *Myc/IgH* translocations murine B cells deficient for *Polq* (Yousefzadeh et al., 2014).

This finding call into question the current paradigm that alternative end joining *in toto* is responsible for the formation of potentially oncogenic translocations. Many oncogenic translocations found in patients contain microhomologies in their breakpoint junctions (Bunting and Nussenzweig, 2013; Weinstock et al., 2006). The subset of broken DNA ends that ultimately give rise to chromosomal translocations are hypothesized to be formed primarily through the microhomology-mediated end joining pathway (MMEJ) subpathway (Bunting and Nussenzweig, 2013). MMEJ relies upon the broken DNA ends annealing at existing microhomologies with only minimal end processing (flap trimming, gap filling, ligation) (Decottignies, 2013). The POLQ-dependent altEJ pathway, which prevents

translocations, is distinct from the MMEJ repair pathway that promotes chromosomal translocations. While no definitive data have been presented to suggest that POLQ has a role in MMEJ, we cannot exclude this possibility. We observed a ten-fold decrease in the repair of substrates with 3' DNA ends containing 4-nucleotide terminal microhomologies in when *Polq* was absent (**Figure 12**) (Yousefzadeh et al., 2014). This finding still allows for a role for POLQ in MMEJ in addition to SD-EJ which is supported by recent biochemical evidence showing that POLQ can perform MMEJ with substrates that contain 3' single-stranded overhangs with at least two nucleotides of homology (Kent et al., 2015).

Why chromosomal translocations form in the absence of POLQ is still unclear. Translocations could form because in the absence of POLQ: (1) DNA breaks persist and eventually free 3' DNA ends find each other and anneal at a microhomology to create a translocation or (2) there is a class of DNA ends that is preferentially handled by POLQ-dependent SD-EJ and are those that form translocations in the absence of POLQ. Further analysis of translocations as proposed in Chapter 3 of this thesis will help clarify how these translocations are formed. A recent study, reported a decrease in insertions at an artificial translocation generated by CRISPR/Cas9 in *Polq*-null MEFs, however the overall frequency of translocations was decreased in the absence of *Polq (Mateos-Gomez et al., 2015)*. It would be interesting to know how POLQ status affects translocations at naturally occurring breaks versus those that are generated through site-specific nucleases.

NGS studies of translocations that occur in mice lacking *Tp53* and *Polq* may demonstrate preferential genomic sites, secondary structures, and sequence specificity that facilitates *Myc/IgH* translocations. Human knockout cell lines, currently being developed by Kei-ichi Takata, could provide insight into the apparent dichotomy that altEJ drives for *in vivo* translocations in mice, whereas, cNHEJ drives translocations in human cell lines (Ghezraoui et al., 2014).

By understanding which proteins POLQ interacts with, we may gain a better understanding as to how POLQ operates in SD-EJ and the repair of DSBs. So far POLQ has only been shown to interact with the replication origin proteins ORC2 and ORC4 (Ramiro et al., 2007). Preliminary data from our collaborator, Dr. Sankar Mitra (Houston Methodist Research Institute), suggests that POLQ interacts with the DNA repair protein XRCC1. XRCC1 is known to interact with DNA Ligase III and DNA polymerase β (Rice, 1999), and the XRCC1-pol β interaction is regulated by HSP90 (Fang et al., 2014). Therefore further investigation of the putative POLQ-XRCC1 interaction may also lead to discovery that HSP90 also interact interacts with POLQ. A proteomics approach to understand the role of POLQ in altEJ is well warranted. Many of the other key proteins involved in altEJ (MRN complex, PARP, XRCC1, DNA Ligases I and III) are also involved in other DNA repair pathways and thus make genetic analysis more difficult to tease out the effects on altEJ (Decottignies, 2013). Phenotypes of certain genetic crosses with *Polq* (ex. XRCC6) could be problematic as loss of both *Atm* and *Polq* causes synthetic lethality in mice. Understanding the detailed mechanism of alternative end joining pathways will give better clarification to how these pathways act and the regulation of alternative end joining with respect to classical non-homologous end joining and HR repair of DNA double-strand breaks.

7.2 POLQ diversifies the immunological portfolio

DNA polymerases play a vital role in adaptive immunity in order to help mammals cope with a multitude of pathogenic insults. Following VDJ recombination, the variable regions of immunoglobulin genes are further diversified through the somatic hypermutation pathway (SHM). SHM relies upon AID-catalyzed deamination of cytosine to generate uracils in the variable region that are processed by UNG to yield abasic sites that are further enzymatically processed. These nucleotide gaps are then filled in by specialized DNA polymerases (Weill and Reynaud, 2008). POLQ was thought to generate most of the A/T mutations found in the variable region of murine B cells, but upon later examination, this class of mutations was largely pol η-dependent (Martomo et al., 2008; Seki et al., 2005). A role for POLQ in SHM may exist whereby POLQ is required to generate SHM-associated indels, as discussed below.

Class switch recombination (CSR) is a programmed break rejoining in the constant region of the immunoglobulin locus. This mechanism allows isotype switching enabling antibodies to interact with different effector molecules without modifying antibody specificity (Bunting and Nussenzweig, 2013; Weill and Reynaud, 2008). Investigation of POLQ function in CSR showed that although cells lacking POLQ continue to switch at a normal frequency compared to their wild-type counterparts, the junctions are different (Martomo et al., 2008; Yousefzadeh et al., 2014). Sequencing of the Sμ-Sγ1 junctions showed that up to 10% of CSR junctions in wild-type mice contained insertions longer than a single nucleotide, and these insertions were dependent upon POLQ (**Figure 11**) (Yousefzadeh et al., 2014). Such insertions form in an altEJ dependent manner (Callen et al., 2009). Our findings support that these insertions arise from POLQ-dependent SD-EJ (Yousefzadeh et al., 2014). Consistent with prior findings, we identified also homologies to nearby sequences around the break in these longer insertions (**Table 4**) (Callen et al., 2009; Yousefzadeh et al., 2014). At least for

CSR junctions, POLQ is responsible for creating the inserted sequences but it may be creating insertions elsewhere in immunoglobulin genes.

Insertions and deletions (indels) that form as a result of SHM, termed SHM-associated (SHA-) indels, represent an understudied but important mechanism of immunological diversity. While most of the purposeful mutations that arise in SHM are point mutations, indels are observed in variable regions at a low frequency (2-6%) (Briney et al., 2012; Wilson et al., 1998). These indels, which are created through AID-mediated breaks, are thought to be a radical new way in which to further evolve our antibody repertoire (Bowers et al., 2014). Currently, it is unknown which polymerase(s) induce these indels in the V regions. However, given that long insertions found in CSR junctions are POLQ-dependent (Yousefzadeh et al., 2014) given evidence that POLQ can cause frameshifts *in vitro* (Arana et al., 2008), we hypothesize that POLQ is the DNA polymerase responsible for the creation these SHA-indels.

SHA-indels are commonly found in the complementarity determining region of antibodies. This is the region where the antibody interacts with the antigen. On rare occasions these SHA-indels are found in the antibody framework regions, which act as the structural scaffold for the antibody (Klein et al., 2013). Indels in the framework regions are generally poorly tolerated, but in rare instances they give rise to unique antibody abilities. One such unique ability is the creation of broadly neutralizing antibodies. Broadly neutralizing antibodies that feature nucleotide insertions in their framework regions are responsible for highly effective responses to pathogens like influenza, HIV-1, and *S. pneumonia* (Krause et al., 2011; Reason and Zhou, 2006; Yu et al., 2008). These framework indels are thought to create new and unconventional ways for antibodies to interact with antigens and notably, deletion of these indels ablates their broadly neutralizing capacity (Klein et al., 2013). Whether POLQ is involved in this rare but distinct mechanism of increasing immunological diversity remains to be seen. The NGS experiments proposed in

Chapter 3 will investigate the whether POLQ regulates the formation of SHA-indels, including those that occur in the framework region in B cells proficient. If POLQ does contribute to the creation of these indels in either the CDR or framework regions, then the regulation of POLQ's activity will be of great interest in developing new immunotherapeutics.

7.3 The impact of POLQ on breast cancer and chemotherapeutics

Reports that depletion or loss of POLQ in murine and human cancer cell lines caused radiosensitivity, stimulated the idea of POLQ as a potential drug target (Begg, 2010; Goff et al., 2009; Higgins et al., 2010b). Upon investigation of replicative stress in breast cancer, our lab and collaborators reported that *POLQ* is the only DNA polymerase that is significantly overexpressed in breast cancer. Further, *POLQ* overexpression is correlated with poorer patient outcomes (Lemée et al., 2010). Subsequent reports also show that overexpression of *POLQ* is associated poorer survival rates in early breast cancers. In addition they identified specific *POLQ* SNPs that were associated with hereditary breast cancer (Brandalize et al., 2014; Higgins et al., 2010a). Two questions float in the ether unanswered right now: (1) If elevated *POLQ* expression leads to poorer outcomes in breast cancer patients, what is the underlying mechanism and (2) if POLQ is of therapeutic value, are there particular breast cancer subtypes that would display a greater therapeutic response when POLQ is inhibited?

I sought to answer the first question of how POLQ causes poorer outcomes by *in vitro* analysis. Preliminary evidence that I present in Chapter 4 demonstrates that ectopic expression of POLQ in 293T cells increases resistance to the DNA damaging agent bleomycin (**Figure 24**). This thesis presents evidence to demonstrate a role for mammalian POLQ in an altEJ repair pathway that is involved in DSB restitution (Yousefzadeh et al., 2014). I hypothesize that this bleomycin resistance seen in 293T cells expressing recombinant POLQ is through its SD-EJ activity. This is supported by the fact that no DNA damage resistance was seen in cells transfected with a polymerase-dead version of POLQ and I have shown that the polymerase activity of POLQ is necessary for its DNA damage resistance activities (Yousefzadeh et al., 2014). Further analysis is needed to see if this DNA damage resistance effect upon ectopic expression of *POLQ* extends to breast cancer cell lines and if increased SD-EJ is occurring in these cells. If ectopic *POLQ* expression

leads to an increase in DNA damage resistance in cell culture models of breast cancer, then further analysis using *in vivo* transgenic mouse models of POLQ would be warranted. I would recommend the use of a conditional transgenic (such as STOP-floxed POLQ and MMTV-Cre) to combat potential issues that may arise from constitutive gain of function of POLQ.

With regard to the second question of whether particular breast cancer subtypes are more responsive to POLQ inhibition, the Wood Laboratory (with the help of C. Marcelo Aldaz) performed a meta-analysis of breast cancer microarray expression data. We found POLQ to be highly expressed in triple negative breast cancers (TNBC), which are refractive to common therapies (unpublished data, C. Marcelo Aldaz and Richard Wood). We hypothesize that by suppressing POLQ by shRNA in TNBC lines we would be able to achieve a significant increase in radiosensitivity in these cells. We are also surveying breast cancer cell lines with a variety of different cancer genotypes (ER, PR, Her2/neu, p53 status) to see if suppression of POLQ would induce hypersensitivity to ionizing radiation (IR). In the mean time, we are investigating small molecule inhibitors of POLQ in the hope that if POLQ-suppression mediates hypersensitivity to IR, that a POLQ-specific inhibitor will cause a similar outcome.

If inhibition of POLQ induces further radiosensitization of breast cancer cells this creates a fantastic therapeutic opportunity. Suppression of POLQ had no effect on a normal human cell line (Higgins et al., 2010b), suggesting only limited off target effects in healthy tissue. Thus delivery (oral or intravenous) of a POLQ-specific inhibitor to patients would make radiotherapy more effective and allow for use of lower doses of radiation. *POLQ* was shown to have as much predictive value for survival outcome for breast patients as *CyclinE* (Lemée et al., 2010), which is currently being used for clinical prognoses (Keyomarsi et al., 2002). Use of a POLQ antibody (like those presented in Chapter 6) could be useful to determine which patient prognosis and which tumor types would be more responsive to

POLQ inhibition. To date no POLQ antibody exists to serve that role but we have many candidate antibodies that require further validation. While much work remains to investigate the role of POLQ in breast cancer outcomes, POLQ presents a novel biomarker that is of diagnostic and therapeutic interest.

7.4 An "origin"al role for POLQ in DNA replication

A recent report demonstrated a role for POLQ in general DNA replication and that loss of POLQ altered the firing of replication origin in cells (Fernandez-Vidal et al., 2014). This finding, coupled with the data that POLQ interacts with the replication origin proteins ORC2 and ORC4, serve as the first suggestions that POLQ operates at the replication fork (Fernandez-Vidal et al., 2014). Expression of *POLQ* correlates with the expression of other genes involved in DNA replication and in regulation of the cell cycle (Pillaire et al., 2009). Ectopic expression of POLQ induces DNA damage signaling, increased chromosomal aberrations, and defective progression of the replication fork (Lemée et al., 2010). In *Arabidopsis*, mutants of the POLQ homolog, TEBICHI, exhibit impairments in meristem cell division and differentiation (Inagaki et al., 2009; Inagaki et al., 2006). Furthermore, a recent study suggested that in human cells POLQ promotes translesion DNA synthesis through thymine glycol adduct (Yoon et al., 2014). These finding add further complexity to the function of POLQ in the cell and now when working with POLQ, we must consider its activities in replication, repair, and damage tolerance. While these *in vitro* findings give pause for thought, these effects do not appear to have a large-scale effect at the organismal level in mice lacking *Polq*. These mice only display a subtle phenotype of increased spontaneous micronuclei (Shima et al., 2004). Therefore, more work is needed to fully understand the biological relevance of POLQ's activities during general DNA replication.

7.5 Conclusion

DNA polymerase θ (POLQ) is a specialized DNA polymerase for which we are only beginning to understand. Studies so far give insight into how this unique enzyme functions in cells, but a considerable amount of information about this enzyme is still unknown. Further exploration of how and with what other proteins POLQ functions with in alternative end joining is required. How the synthesis-dependent end joining pathway, which is predicated upon the polymerase activities of POLQ, functions in relation to other DNA double-strand break pathways is still unclear. How DNA breaks are channeled into this specific pathway in lieu of others is still a mystery. Is the nucleotide insertion activity of POLQ salvaging genomic integrity and preventing *Myc*/*IgH* translocation formation at the cost of creating indels? This is our current hypothesis but additional evidence is needed to substantiate this notion.

The function of POLQ in other activities besides DNA double-strand break repair is intriguing but requires more validation. If POLQ regulates origin firing and replication dynamics, it would be valuable to know to what extent and the biological consequences of misregulation. Also does POLQ share responsibility for these actions with any of the other specialized DNA polymerases remains to be seen. Based on the published literature and preliminary data generated in the lab, there are many unanswered questions about POLQ. POLQ aids in maintaining a healthy genomic relationship through rejoining broken DNA ends by inserting some common ground. We have only a limited knowledge of how it does this, temporospatial nature of this activity, and how to exploit POLQ for better therapies. But there is much to learn about POLQ and the sky is the limit.

Appendix 1. Analysis of REV7 in breast cancer and DNA damage sensitivity.

DNA polymerase ζ, another specialized DNA polymerase that is studied by the lab, is includes REV3L and REV7 that form the catalytic and accessory subunits of pol ζ, respectively. Pol ζ is required for viability and defends cells against chromosomal instability, as well as suppresses spontaneous tumor formation *in vivo* (Lange et al., 2011). However, how human REV7 interacts with REV3L and the biological consequences of this interaction on the human pol ζ complex remains to be clarified. Dr. Junya Tomida of our lab discovered that in addition to a previously identified binging site, REV3L also contains a second REV7 binding motif (Tomida et al., 2015). Further work demonstrated that both REV7 binding sites must be present for pol ζ to confer resistance to DNA damage (unpublished data, Junya Tomida and Richard Wood). Dr. Tomida also found that reduced REV7 protein levels increased the abundance of REV3L and that ectopic expression of *REV7* in HeLa cells induced mitomycin c (MMC) hypersensitivity compared to the empty vector control (unpublished data, Junya Tomida) (Tomida et al., 2015).

Previously, I had surveyed expression of a number of DNA repair genes in breast cancer cell lines (**Figure 19**) and found an elevated level of *REV7* expression in HCC1806 cells (**Figure 19E, Figure 30A**), which is a triple negative breast cancer line that lacks functional p53 (Neve et al., 2006). Following this, I surveyed the HCC1806 (*REV7* high) and MCF7 cell (*REV7* low) lines for MMC sensitivity at different time points (**Figure 30B**). Compared to MCF7 cells, HCC1806 cells display greater sensitivity to MMC, however this may not be dependent upon REV7 levels alone (**Figure 30B**). To test the hypothesis that elevated levels of REV7 confer DNA damage sensitivity, MCF7 (*REV7* low) cells were transiently transfected with a recombinant human *REV7* expression vector or an empty vector control (**Figure 30C and D**). MCF7 cells that expressed recombinant human REV7 were much more hypersensitive to MMC than cells containing the empty vector control plasmid (**Figure 30C**).

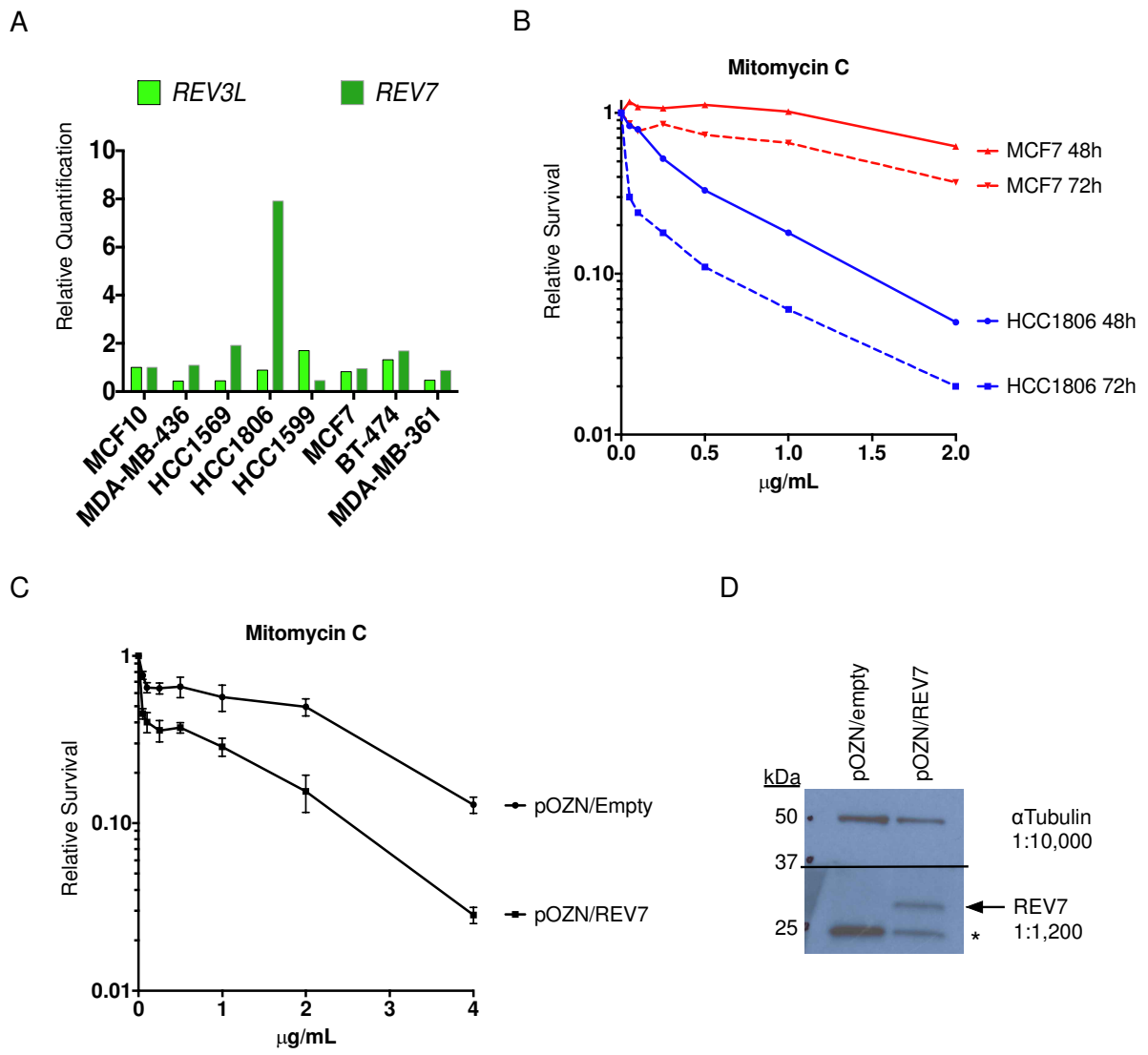


Figure 30. Ectopic expression of REV7 can induce DNA damage hypersensitivity in breast cancer cells. (**A**) Expression of the indicated genes was quantified by qPCR in multiple breast cancer cells. Relative quantification of gene expression was examined using Taqman probes. Data was analyzed using the ΔΔCt method and *GAPDH* expression was used as an internal control. (**B**) MCF7 and HCC1806 breast cancer cells were treated with mitomycin c and surveyed for ATP levels as a surrogate marker of survival at 48 and 72 h after damage. (**C**) Similarly MC7 cells were transiently transfected with either pOZN/empty or pOZN/REV7. Transfected cells were replated and treated with the indicated

concentrations of mitomycin c and ATP levels were similarly measured 48 h after damage. Data reflects three independent experiments. (**D**) Representative immunoblot of pooled crude SDS lysates from the three independent experiments. The asterisk denotes endogenous REV7, while the arrow indicates the recombinant N-terminal FLAG-HA-tagged REV7. Cell lines were cultured and maintained by MJY. MJY isolated RNA and analyzed qPCR expression data that was gathered by the Molecular Biology Core.

Further analysis of these preliminary findings is warranted. Analysis of breast cancer outcomes and disease-free survival correlated to *REV3L* and *REV7* levels or analysis of tumor samples by REV7 immunohistochemistry is needed. The finding that ectopic expression of *REV7* increases DNA damage sensitivity in cultured cells should be further evaluated by examining increased and decreased levels of REV7 protein in *REV7* high expressing cell lines, like HCC1806 (**Figure 30**). If elevated REV7 abundance is found in these cells then ectopic expression of recombinant REV7 should have minimal effect protein levels are already at a level that is saturating. However, once protein levels reach a certain point, it is plausible that protein could become cytotoxic. Also REV7 levels are tightly regulated and shRNA suppression of *REV7* could rescue the MMC sensitivity phenotype, but only up until a point as *REV7*-deficient chicken DT40 cells are hypersensitive to MMC and cisplatin (Okada et al., 2005). Thus there is a fine balance required in regard to the regulation of REV7 levels as too much or too little could pose grave consequences to the cell.

Material and methods

Please note that this section is written as a protocol for persons that may be trying to repeat these experimental procedures.

Cell lines

The following cell lines were purchased from American Type Culture Collection: MCF7, MCF10, HCC1569, HCC1599, HCC1806, BT-474, MDA-MB-361, and MDA-MB-436. Cells were STR fingerprinted, checked for mycoplasma contamination, and cultured in their appropriate media (Neve et al., 2006).

RNA isolation and qPCR analysis

Total RNA was isolated from the each of the breast cancer cell lines were analyzed for quality and purity using RNA 6000 Nano kit (Agilent Technologies, Santa Clara, CA). 1 µg of total RNA was used to generate cDNA using the High Capacity cDNA RT kit (Life Technologies). qPCR analysis was performed in triplicate using the ABI Prism 7900 HT thermocycler and the following Taqman Probe set (Bio-Rad, Hercules, CA): *REV3L* (Hs01076848_m1); *REV7* (Hs01057448_m1). Data were analyzed using the ΔΔCt method. *GAPDH* was used as an internal control in all experiments and relative quantification analysis was performed with MCF10 cells serving as the non-tumorigenic control.

Transfection of 293T cells and Crude SDS Lysate Preparation

1.5 x10^5 293T cells were plated into each well of a 6-well dish. The following day, cells were transfected with 2 µg of either pOZN/empty vector control or pOZN/REV7 that was provided by Dr. Junya Tomida. 48 h post-transfection cells from one well of a 6 well plate were resuspended in 200 µL of 2X SDS Loading Buffer. Lysates were heat denatured at 95°C for 5 min.

2X SDS Loading Buffer

20% Glycerol

4% SDS

0.2% Bromophenol Blue

12% 2-Mercaptoethanol

100 mM Tris-HCl pH 6.8

Appendix 2. Analysis of *Fam35a* expression in murine tissues.

FAM35A was identified by co-immunoprecipitation/mass spectrometry analysis as an interaction partner of human REV7, the catalytic subunit of DNA polymerase ζ (unpublished data, Junya Tomida and Richard Wood). Little is known about the protein, but suppression of FAM35A in human cancer cell lines induces hypersensitivity to the interstrand-crosslinking agent mitomycin c and causes elevated levels of DNA damage based on enumeration of RAD51 and γH2AX foci (unpublished data, Junya Tomida and Richard Wood). To get a better understanding of FAM35A and its role in cells, I surveyed *Fam35a* transcript levels in murine tissues from C57BL/6J mice by qPCR using absolute quantification method (**Figure 31**).

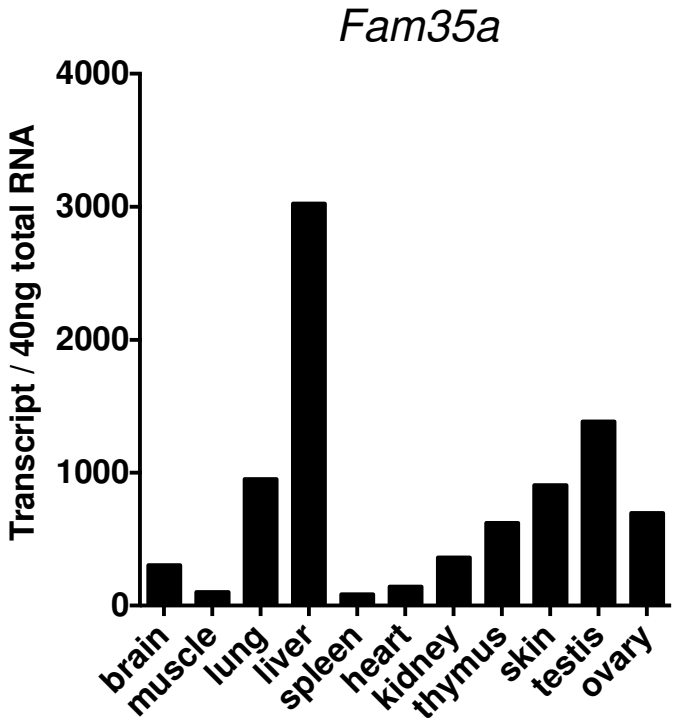


Figure 31. Absolute quantification of transcript number shows that *Fam35a* is broadly expressed in tissues. Total RNA was isolated from necropsied C57BL/6J mice by Trizol (Life Technologies). Transcript abundance for *Fam35a* was determined by absolute

quantification method as described in the material and methods section. *Gapdh* was used as an internal control to normalize samples. Animal tissue was harvested and isolated for RNA by MJY. MJY analyzed qPCR expression data generated by the Science Park Molecular Biology Core.

Fam35a is broadly expressed in a number of tissues from mice, with the highest expression in the liver. Further analysis (genetic, proteomic, etc.) of this gene in mammals will reveal its cellular function and which DNA repair pathways that FAM35A could participate in.

Material and methods

Quantitative real-time PCR analysis of mouse tissues. RNA was isolated from tissues of C57BL/6J mice by Trizol and was analyzed for quality and purity using a RNA 6000 Nano kit (Agilent Technologies, Santa Clara, CA). 1 µg of total RNA was used to generate cDNA using the High Capacity cDNA RT kit (Life Technologies). qPCR analysis was performed in triplicate using the ABI Prism 7900 HT thermocycler and the *Fam35a* (Mm01209709_m1) Taqman Probe and primer set with iTAQ SYBR Green Supermix with ROX (Bio-Rad, Hercules, CA). Data were analyzed using the absolute quantification method, where titration of pT7T3D-PacI-Fam35a-AA242674 plasmid was used to generate standard curves for expression. Transcript abundance was determined by extrapolation from linear regression analysis of best-fit lines from titration experiments. *Gapdh* was used as an internal control for normalization.

BIBLIOGRAPHY

Abbas, A. (2000). Cloning and chromosomal mapping of the human DNA polymerase theta (University of California at Berkeley).

Aguirrezabalaga, I., Nivard, M.J., Comendador, M.A., and Vogel, E.W. (1995). The cross-linking agent hexamethylphosphoramide predominantly induces intra-locus and multi-locus deletions in postmeiotic germ cells of Drosophila. Genetics *139*, 649-658.

André, F., and Zielinski, C.C. (2012). Optimal strategies for the treatment of metastatic triple-negative breast cancer with currently approved agents. Ann Oncol *23 Suppl 6*, vi46-51.

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

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

Asagoshi, K., Lehmann, W., Braithwaite, E.K., Santana-Santos, L., Prasad, R., Freedman, J.H., Van Houten, B., and Wilson, S.H. (2012). Single-nucleotide base excision repair DNA polymerase activity in C. elegans in the absence of DNA polymerase beta. Nucleic Acids Res *40*, 670-681.

Avkin, S., Adar, S., Blander, G., and Livneh, Z. (2002). Quantitative measurement of translesion replication in human cells: evidence for bypass of abasic sites by a replicative DNA polymerase. Proc Natl Acad Sci U S A *99*, 3764-3769.

Begg, A. (2010). POLQ in breast cancer. Oncotarget *1*, 161-162.

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

Bhagwat, N., Olsen, A.L., Wang, A.T., Hanada, K., Stuckert, P., Kanaar, R., D'Andrea, A., Niedernhofer, L.J., and McHugh, P.J. (2009). XPF-ERCC1 participates in the Fanconi anemia pathway of cross-link repair. Mol Cell Biol *29*, 6427-6437.

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

Boboila, C., Oksenych, V., Gostissa, M., Wang, J.H., Zha, S., Zhang, Y., Chai, H., Lee, C.S., Jankovic, M., Saez, L.M., Nussenzweig, M.C., McKinnon, P.J., Alt, F.W., and Schwer, B. (2012b). Robust chromosomal DNA repair via alternative end-joining in the absence of X-ray repair cross-complementing protein 1 (XRCC1). Proc Natl Acad Sci U S A *109*, 2473-2478.

Bowers, P.M., Verdino, P., Wang, Z., da Silva Correia, J., Chhoa, M., Macondray, G., Do, M., Neben, T.Y., Horlick, R.A., Stanfield, R.L., Wilson, I.A., and King, D.J. (2014). Nucleotide insertions and deletions complement point mutations to massively expand the diversity created by somatic hypermutation of antibodies. J Biol Chem *289*, 33557-33567.

Brandalize, A.P., Schuler-Faccini, L., Hoffmann, J.S., Caleffi, M., Cazaux, C., and Ashton-Prolla, P. (2014). A DNA repair variant in POLQ (c.-1060A > G) is associated to hereditary breast cancer patients: a case-control study. BMC Cancer *14*, 850.

Briney, B.S., Willis, J.R., and Crowe, J.E., Jr. (2012). Location and length distribution of somatic hypermutation-associated DNA insertions and deletions reveals regions of antibody structural plasticity. Genes Immun *13*, 523-529.

Brosh, R.M., Jr., and Cantor, S.B. (2014). Molecular and cellular functions of the FANCJ DNA helicase defective in cancer and in Fanconi anemia. Front Genet *5*, 372.

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

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

Callen, E., Jankovic, M., Wong, N., Zha, S., Chen, H.T., Difilippantonio, S., Di Virgilio, M., Heidkamp, G., Alt, F.W., Nussenzweig, A., and Nussenzweig, M. (2009). Essential role for DNA-PKcs in DNA double-strand break repair and apoptosis in ATM-deficient lymphocytes. Mol Cell *34*, 285-297.

Cannistraro, V.J., and Taylor, J.S. (2004). DNA-thumb interactions and processivity of T7 DNA polymerase in comparison to yeast polymerase eta. J Biol Chem *279*, 18288-18295.

Ceccaldi, R., Liu, J.C., Amunugama, R., Hajdu, I., Primack, B., Petalcorin, M.I.R., O'Connor, K.W., Konstantinopoulos, P.A., Elledge, S.J., Boulton, S.J., Yusufzai, T., and D'Andrea, A. (2015). Homologous-recombination-deficient tumours are dependent on Pol theta-mediated repair. Nature.

Chahwan, R., Edelmann, W., Scharff, M.D., and Roa, S. (2011). Mismatch-mediated error prone repair at the immunoglobulin genes. Biomed Pharmacother *65*, 529-536.

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

Cheng, C.H., and Kuchta, R.D. (1993). DNA polymerase epsilon: aphidicolin inhibition and the relationship between polymerase and exonuclease activity. Biochemistry *32*, 8568-8574.

Chiarle, R., Zhang, Y., Frock, R.L., Lewis, S.M., Molinie, B., Ho, Y.J., Myers, D.R., Choi, V.W., Compagno, M., Malkin, D.J., Neuberg, D., Monti, S., Giallourakis, C.C., Gostissa, M., and Alt, F.W. (2011). Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. Cell *147*, 107-119.

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

Deriano, L., and Roth, D.B. (2013). Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. Annu Rev Genet *47*, 433-455.

Dertinger, S.D., Torous, D.K., and Tometsko, K.R. (1996). Simple and reliable enumeration of micronucleated reticulocytes with a single-laser flow cytometer. Mutat Res *371*, 283-292.

Di Noia, J.M., and Neuberger, M.S. (2007). Molecular mechanisms of antibody somatic hypermutation. Annu Rev Biochem *76*, 1-22.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res *11*, 1475-1489.

Dorsett, Y., McBride, K.M., Jankovic, M., Gazumyan, A., Thai, T.H., Robbiani, D.F., Di Virgilio, M., Reina San-Martin, B., Heidkamp, G., Schwickert, T.A., Eisenreich, T., Rajewsky, K., and Nussenzweig, M.C. (2008). MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. Immunity *28*, 630-638.

Doublié, S., Tabor, S., Long, A.M., Richardson, C.C., and Ellenberger, T. (1998). Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 A resolution. Nature *391*, 251-258.

Dresler, S.L., Gowans, B.J., Robinson-Hill, R.M., and Hunting, D.J. (1988). Involvement of DNA polymerase delta in DNA repair synthesis in human fibroblasts at late times after ultraviolet irradiation. Biochemistry *27*, 6379-6383.

Fang, Q., Inanc, B., Schamus, S., Wang, X.H., Wei, L., Brown, A.R., Svilar, D., Sugrue, K.F., Goellner, E.M., Zeng, X., Yates, N.A., Lan, L., Vens, C., and Sobol, R.W. (2014). HSP90 regulates DNA repair via the interaction between XRCC1 and DNA polymerase beta. Nat Commun *5*, 5513.

Farmer, H., McCabe, N., Lord, C.J., Tutt, A.N., Johnson, D.A., Richardson, T.B., Santarosa, M., Dillon, K.J., Hickson, I., Knights, C., Martin, N.M., Jackson, S.P., Smith, G.C., and Ashworth, A. (2005). Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature *434*, 917-921.

Ferguson, D.O., Sekiguchi, J.M., Chang, S., Frank, K.M., Gao, Y., DePinho, R.A., and Alt, F.W. (2000). The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations. Proc Natl Acad Sci U S A *97*, 6630-6633.

Fernandez-Vidal, A., Guitton-Sert, L., Cadoret, J.C., Drac, M., Schwob, E., Baldacci, G., Cazaux, C., and Hoffmann, J.S. (2014). A role for DNA polymerase theta in the timing of DNA replication. Nat Commun *5*, 4285.

Frank-Vaillant, M., and Marcand, S. (2002). Transient stability of DNA ends allows nonhomologous end joining to precede homologous recombination. Mol Cell *10*, 1189-1199.

Friedberg, E.C., Walker, G.C., Siede, W., Wood, R.D., Schultz, R.A., and Ellenberger, T. (2006). DNA Repair and Mutagenesis, 2nd edition edn (Washington, DC: ASM Press).

Frit, P., Barboule, N., Yuan, Y., Gomez, D., and Calsou, P. (2014). Alternative end-joining pathway(s): bricolage at DNA breaks. DNA Repair *17*, 81-97.

Garcia-Gomez, S., Reyes, A., Martinez-Jimenez, M.I., Chocron, E.S., Mouron, S., Terrados, G., Powell, C., Salido, E., Mendez, J., Holt, I.J., and Blanco, L. (2013). PrimPol, an archaic primase/polymerase operating in human cells. Mol Cell *52*, 541-553.

Gazumyan, A., Bothmer, A., Klein, I.A., Nussenzweig, M.C., and McBride, K.M. (2012). Activation-induced cytidine deaminase in antibody diversification and chromosome translocation. Adv Cancer Res *113*, 167-190.

Ghezraoui, H., Piganeau, M., Renouf, B., Renaud, J.B., Sallmyr, A., Ruis, B., Oh, S., Tomkinson, A.E., Hendrickson, E.A., Giovannangeli, C., Jasin, M., and Brunet, E. (2014). Chromosomal translocations in human cells are generated by canonical nonhomologous end-joining. Mol Cell *55*, 829-842.

Goff, J.P., Shields, D.S., Seki, M., Choi, S., Epperly, M.W., Dixon, T., Wang, H., Bakkenist, C.J., Dertinger, S.D., Torous, D.K., Wittschieben, J., Wood, R.D., and Greenberger, J.S. (2009). Lack of DNA polymerase theta (POLQ) radiosensitizes bone marrow stromal cells in vitro and increases reticulocyte micronuclei after total-body irradiation. Radiat Res *172*, 165-174.

Goodman, M.F. (2002). Error-prone repair DNA polymerases in prokaryotes and eukaryotes. Annu Rev Biochem *71*, 17-50.

Guillet, M., and Boiteux, S. (2003). Origin of endogenous DNA abasic sites in Saccharomyces cerevisiae. Mol Cell Biol *23*, 8386-8394.

Hamer, G., Roepers-Gajadien, H.L., van Duyn-Goedhart, A., Gademan, I.S., Kal, H.B., van Buul, P.P., and de Rooij, D.G. (2003). DNA double-strand breaks and gamma-H2AX signaling in the testis. Biol Reprod *68*, 628-634.

Han, L., Masani, S., Hsieh, C.L., and Yu, K. (2014). DNA ligase I is not essential for mammalian cell viability. Cell Rep *7*, 316-320.

Harrigan, J.A., Belotserkovskaya, R., Coates, J., Dimitrova, D.S., Polo, S.E., Bradshaw, C.R., Fraser, P., and Jackson, S.P. (2011). Replication stress induces 53BP1-containing OPT domains in G1 cells. J Cell Biol *193*, 97-108.

Harris, P.V., Mazina, O.M., Leonhardt, E.A., Case, R.B., Boyd, J.B., and Burtis, K.C. (1996). 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.

Helmink, B.A., and Sleckman, B.P. (2012). The response to and repair of RAG-mediated DNA double-strand breaks. Annu Rev Immunol *30*, 175-202.

Higgins, G.S., Harris, A.L., Prevo, R., Helleday, T., McKenna, W.G., and Buffa, F.M. (2010a). Overexpression of POLQ confers a poor prognosis in early breast cancer patients. Oncotarget *1*, 175-184.

Higgins, G.S., Prevo, R., Lee, Y.F., Helleday, T., Muschel, R.J., Taylor, S., Yoshimura, M., Hickson, I.D., Bernhard, E.J., and McKenna, W.G. (2010b). A small interfering RNA screen of genes involved in DNA repair identifies tumor-specific radiosensitization by POLQ knockdown. Cancer Res *70*, 2984-2993.

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

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

Honjo, T., Muramatsu, M., and Fagarasan, S. (2004). AID: how does it aid antibody diversity? Immunity *20*, 659-668.

Huang, K.C., Gao, H., Yamasaki, E.F., Grabowski, D.R., Liu, S., Shen, L.L., Chan, K.K., Ganapathi, R., and Snapka, R.M. (2001). Topoisomerase II poisoning by ICRF-193. J Biol Chem *276*, 44488-44494.

Hubscher, U., Maga, G., and Spadari, S. (2002). Eukaryotic DNA polymerases. Annu Rev Biochem *71*, 133-163.

Inagaki, S., Nakamura, K., and Morikami, A. (2009). A link among DNA replication, recombination, and gene expression revealed by genetic and genomic analysis of TEBICHI gene of Arabidopsis thaliana. PLOS Genet *5*, e1000613.

Inagaki, S., Suzuki, T., Ohto, M.A., Urawa, H., Horiuchi, T., Nakamura, K., and Morikami, A. (2006). Arabidopsis TEBICHI, with helicase and DNA polymerase domains, is required for regulated cell division and differentiation in meristems. Plant Cell *18*, 879-892.

Johnson, R.E., Washington, M.T., Haracska, L., Prakash, S., and Prakash, L. (2000). Eukaryotic polymerases iota and zeta act sequentially to bypass DNA lesions. Nature *406*, 1015-1019.

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

Kawamura, K., Bahar, R., Seimiya, M., Chiyo, M., Wada, A., Okada, S., Hatano, M., Tokuhisa, T., Kimura, H., Watanabe, S., Honda, I., Sakiyama, S., Tagawa, M., and J, O.W. (2004). DNA polymerase theta is preferentially expressed in lymphoid tissues and upregulated in human cancers. Int J Cancer *109*, 9-16.

Kelley, M. (2011). DNA Repair in Cancer Therapy (London, UK: Academic Press).

Kent, T., Chandramouly, G., McDevitt, S.M., Ozdemir, A.Y., and Pomerantz, R.T. (2015). Mechanism of microhomology-mediated end-joining promoted by human polymerase theta. Nat Struct Mol Biol.

Keyomarsi, K., Tucker, S.L., Buchholz, T.A., Callister, M., Ding, Y., Hortobagyi, G.N., Bedrosian, I., Knickerbocker, C., Toyofuku, W., Lowe, M., Herliczek, T.W., and Bacus, S.S. (2002). Cyclin E and survival in patients with breast cancer. N Engl J Med *347*, 1566-1575.

Klein, F., Diskin, R., Scheid, J.F., Gaebler, C., Mouquet, H., Georgiev, I.S., Pancera, M., Zhou, T., Incesu, R.B., Fu, B.Z., Gnanapragasam, P.N., Oliveira, T.Y., Seaman, M.S., Kwong, P.D., Bjorkman, P.J., and Nussenzweig, M.C. (2013). Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. Cell *153*, 126-138.

Klein, I.A., Resch, W., Jankovic, M., Oliveira, T., Yamane, A., Nakahashi, H., Di Virgilio, M., Bothmer, A., Nussenzweig, A., Robbiani, D.F., Casellas, R., and Nussenzweig, M.C. (2011). Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. Cell *147*, 95-106.

Kohzaki, M., Nishihara, K., Hirota, K., Sonoda, E., Yoshimura, M., Ekino, S., Butler, J.E., Watanabe, M., Halazonetis, T.D., and Takeda, S. (2010). DNA polymerases nu and theta are required for efficient immunoglobulin V gene diversification in chicken. J Cell Biol *189*, 1117-1127.

Koole, W., van Schendel, R., Karambelas, A.E., van Heteren, J.T., Okihara, K.L., and Tijsterman, M. (2014). A Polymerase Theta-dependent repair pathway suppresses extensive genomic instability at endogenous G4 DNA sites. Nat Commun *5*, 3216.

Kottemann, M.C., and Smogorzewska, A. (2013). Fanconi anaemia and the repair of Watson and Crick DNA crosslinks. Nature *493*, 356-363.

Kovalchuk, A.L., Muller, J.R., and Janz, S. (1997). Deletional remodeling of c-myc-deregulating chromosomal translocations. Oncogene *15*, 2369-2377.

Krause, J.C., Ekiert, D.C., Tumpey, T.M., Smith, P.B., Wilson, I.A., and Crowe, J.E., Jr. (2011). An insertion mutation that distorts antibody binding site architecture enhances function of a human antibody. mBio *2*, e00345-00310.

Kusumoto, R., Masutani, C., Iwai, S., and Hanaoka, F. (2002). Translesion synthesis by human DNA polymerase eta across thymine glycol lesions. Biochemistry *41*, 6090-6099.

Lange, S.S., Takata, K., and Wood, R.D. (2011). DNA polymerases and cancer. Nat Rev Cancer *11*, 96-110.

Lange, S.S., Wittschieben, J.P., and Wood, R.D. (2012). DNA polymerase zeta is required for proliferation of normal mammalian cells. Nucleic Acids Res *40*, 4473-4482.

Lawrence, C.W. (2002). Cellular roles of DNA polymerase zeta and Rev1 protein. DNA Repair *1*, 425-435.

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

Lee, K.A., Bindereif, A., and Green, M.R. (1988). A small-scale procedure for preparation of nuclear extracts that support efficient transcription and pre-mRNA splicing. Gene Anal Tech *5*, 22-31.

Lehmann, B.D., Bauer, J.A., Chen, X., Sanders, M.E., Chakravarthy, A.B., Shyr, Y., and Pietenpol, J.A. (2011). Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest *121*, 2750-2767.

Lemée, F., Bergoglio, V., Fernandez-Vidal, A., Machado-Silva, A., Pillaire, M.J., Bieth, A., Gentil, C., Baker, L., Martin, A.L., Leduc, C., Lam, E., Magdeleine, E., Filleron, T., Oumouhou, N., Kaina, B., Seki, M., Grimal, F., Lacroix-Triki, M., Thompson, A., Roche, H., Bourdon, J.C., Wood, R.D., Hoffmann, J.S., and Cazaux, C. (2010). 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.

Li, Y., Gao, X., and Wang, J.Y. (2011). Comparison of two POLQ mutants reveals that a polymerase-inactive POLQ retains significant function in tolerance to etoposide and gamma-irradiation in mouse B cells. Genes Cells *16*, 973-983.

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

Loveday, C., Turnbull, C., Ramsay, E., Hughes, D., Ruark, E., Frankum, J.R., Bowden, G., Kalmyrzaev, B., Warren-Perry, M., Snape, K., Adlard, J.W., Barwell, J., Berg, J., Brady, A.F., Brewer, C., Brice, G., Chapman, C., Cook, J., Davidson, R., Donaldson, A., Douglas, F., Greenhalgh, L., Henderson, A., Izatt, L., Kumar, A., Lalloo, F., Miedzybrodzka, Z., Morrison, P.J., Paterson, J., Porteous, M., Rogers, M.T., Shanley, S., Walker, L., Breast Cancer Susceptibility, C., Eccles, D., Evans, D.G., Renwick, A., Seal, S., Lord, C.J., Ashworth, A., Reis-Filho, J.S., Antoniou, A.C., and Rahman, N. (2011). Germline mutations in RAD51D confer susceptibility to ovarian cancer. Nat Genet *43*, 879-882.

Lukas, C., Savic, V., Bekker-Jensen, S., Doil, C., Neumann, B., Pedersen, R.S., Grofte, M., Chan, K.L., Hickson, I.D., Bartek, J., and Lukas, J. (2011). 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. Nat Cell Biol *13*, 243-253.

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

Mahadevaiah, S.K., Turner, J.M., Baudat, F., Rogakou, E.P., de Boer, P., Blanco-Rodriguez, J., Jasin, M., Keeney, S., Bonner, W.M., and Burgoyne, P.S. (2001). Recombinational DNA double-strand breaks in mice precede synapsis. Nat Genet *27*, 271-276.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., and Bryant, S.H. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res *39*, D225-229.

Marini, F., Kim, N., Schuffert, A., and Wood, R.D. (2003). POLN, a nuclear PolA family DNA polymerase homologous to the DNA cross-link sensitivity protein Mus308. J Biol Chem *278*, 32014-32019.

Marini, F., and Wood, R.D. (2002). A human DNA helicase homologous to the DNA cross-link sensitivity protein Mus308. J Biol Chem *277*, 8716-8723.

Marti, T.M., Hefner, E., Feeney, L., Natale, V., and Cleaver, J.E. (2006). H2AX phosphorylation within the G1 phase after UV irradiation depends on nucleotide excision repair and not DNA double-strand breaks. Proc Natl Acad Sci U S A *103*, 9891-9896.

Martomo, S.A., Saribasak, H., Yokoi, M., Hanaoka, F., and Gearhart, P.J. (2008). Reevaluation of the role of DNA polymerase theta in somatic hypermutation of immunoglobulin genes. DNA Repair *7*, 1603-1608.

Masuda, K., Ouchida, R., Hikida, M., Kurosaki, T., Yokoi, M., Masutani, C., Seki, M., Wood, R.D., Hanaoka, F., and J, O.W. (2007). DNA polymerases eta and theta function in the same genetic pathway to generate mutations at A/T during somatic hypermutation of Ig genes. J Biol Chem *282*, 17387-17394.

Masuda, K., Ouchida, R., Takeuchi, A., Saito, T., Koseki, H., Kawamura, K., Tagawa, M., Tokuhisa, T., Azuma, T., and J, O.W. (2005). DNA polymerase theta contributes to the generation of C/G mutations during somatic hypermutation of Ig genes. Proc Natl Acad Sci U S A *102*, 13986-13991.

Masutani, C., Kusumoto, R., Iwai, S., and Hanaoka, F. (2000). Mechanisms of accurate translesion synthesis by human DNA polymerase eta. EMBO J *19*, 3100-3109.

Mateos-Gomez, P.A., Gong, F., Nair, N., Miller, K.M., Lazzerini-Denchi, E., and Sfeir, A. (2015). Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination. Nature *518*, 254-257.

Minnick, D.T., Astatke, M., Joyce, C.M., and Kunkel, T.A. (1996). A thumb subdomain mutant of the large fragment of Escherichia coli DNA polymerase I with reduced DNA binding affinity, processivity, and frameshift fidelity. J Biol Chem *271*, 24954-24961.

Moldovan, G.L., Madhavan, M.V., Mirchandani, K.D., McCaffrey, R.M., Vinciguerra, P., and D'Andrea, A.D. (2010). DNA polymerase POLN participates in cross-link repair and homologous recombination. Mol Cell Biol *30*, 1088-1096.

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

Nakamura, J., Walker, V.E., Upton, P.B., Chiang, S.Y., Kow, Y.W., and Swenberg, J.A. (1998). Highly sensitive apurinic/apyrimidinic site assay can detect spontaneous and chemically induced depurination under physiological conditions. Cancer Res *58*, 222-225.

Neuberger, M.S., Di Noia, J.M., Beale, R.C., Williams, G.T., Yang, Z., and Rada, C. (2005). Somatic hypermutation at A.T pairs: polymerase error versus dUTP incorporation. Nat Rev Immunol *5*, 171-178.

Neve, R.M., Chin, K., Fridlyand, J., Yeh, J., Baehner, F.L., Fevr, T., Clark, L., Bayani, N., Coppe, J.P., Tong, F., Speed, T., Spellman, P.T., DeVries, S., Lapuk, A., Wang, N.J., Kuo, W.L., Stilwell, J.L., Pinkel, D., Albertson, D.G., Waldman, F.M., McCormick, F., Dickson, R.B., Johnson, M.D., Lippman, M., Ethier, S., Gazdar, A., and Gray, J.W. (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell *10*, 515-527.

Nusse, M., Miller, B.M., Viaggi, S., and Grawe, J. (1996). Analysis of the DNA content distribution of micronuclei using flow sorting and fluorescent in situ hybridization with a centromeric DNA probe. Mutagenesis *11*, 405-413.

Okada, T., Sonoda, E., Yoshimura, M., Kawano, Y., Saya, H., Kohzaki, M., and Takeda, S. (2005). Multiple roles of vertebrate REV genes in DNA repair and recombination. Mol Cell Biol *25*, 6103-6111.

Patel, P.H., and Loeb, L.A. (2000). DNA polymerase active site is highly mutable: evolutionary consequences. Proc Natl Acad Sci U S A *97*, 5095-5100.

Pillaire, M.J., Selves, J., Gordien, K., Gouraud, P.A., Gentil, C., Danjoux, M., Do, C., Negre, V., Bieth, A., Guimbaud, R., Trouche, D., Pasero, P., Mechali, M., Hoffmann, J.S., and Cazaux, C. (2009). A 'DNA replication' signature of progression and negative outcome in colorectal cancer. Oncogene.

Plecenikova, A., Slaninova, M., and Riha, K. (2014). Characterization of DNA repair deficient strains of Chlamydomonas reinhardtii generated by insertional mutagenesis. PLOS One *9*, e105482.

Prasad, R., Longley, M.J., Sharief, F.S., Hou, E.W., Copeland, W.C., and Wilson, S.H. (2009). 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.

Rajewsky, K. (1996). Clonal selection and learning in the antibody system. Nature *381*, 751-758.

Ramiro, A., Reina San-Martin, B., McBride, K., Jankovic, M., Barreto, V., Nussenzweig, A., and Nussenzweig, M.C. (2007). The role of activation-induced deaminase in antibody diversification and chromosome translocations. Adv Immunol *94*, 75-107.

Ramiro, A.R., Jankovic, M., Callen, E., Difilippantonio, S., Chen, H.T., McBride, K.M., Eisenreich, T.R., Chen, J., Dickins, R.A., Lowe, S.W., Nussenzweig, A., and Nussenzweig, M.C. (2006). Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. Nature *440*, 105-109.

Ramsden, D.A., and Asagoshi, K. (2012). DNA polymerases in nonhomologous end joining: are there any benefits to standing out from the crowd? Environ Mol Mutagen *53*, 741-751.

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

Reason, D.C., and Zhou, J. (2006). Codon insertion and deletion functions as a somatic diversification mechanism in human antibody repertoires. Biol Direct *1*, 24.

Reina-San-Martin, B., Difilippantonio, S., Hanitsch, L., Masilamani, R.F., Nussenzweig, A., and Nussenzweig, M.C. (2003). H2AX is required for recombination between immunoglobulin switch regions but not for intra-switch region recombination or somatic hypermutation. J Exp Med *197*, 1767-1778.

Rice, P.A. (1999). Holding damaged DNA together. Nat Struct Mol Biol *6*, 805-806.

Robbiani, D.F., Bothmer, A., Callen, E., Reina-San-Martin, B., Dorsett, Y., Difilippantonio, S., Bolland, D.J., Chen, H.T., Corcoran, A.E., Nussenzweig, A., and Nussenzweig, M.C. (2008). AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. Cell *135*, 1028-1038.

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

Sale, J.E., and Neuberger, M.S. (1998). TdT-accessible breaks are scattered over the immunoglobulin V domain in a constitutively hypermutating B cell line. Immunity *9*, 859-869.

Sarkies, P., Murat, P., Phillips, L.G., Patel, K.J., Balasubramanian, S., and Sale, J.E. (2012). FANCJ coordinates two pathways that maintain epigenetic stability at G-quadruplex DNA. Nucleic Acids Res *40*, 1485-1498.

Seki, M., Gearhart, P.J., and Wood, R.D. (2005). DNA polymerases and somatic hypermutation of immunoglobulin genes. EMBO Rep *6*, 1143-1148.

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

Seki, M., Masutani, C., Yang, L.W., Schuffert, A., Iwai, S., Bahar, I., and Wood, R.D. (2004). High-efficiency bypass of DNA damage by human DNA polymerase Q. EMBO J *23*, 4484-4494.

Seki, M., and Wood, R.D. (2008). DNA polymerase theta (POLQ) can extend from mismatches and from bases opposite a (6-4) photoproduct. DNA Repair *7*, 119-127.

Sharief, F.S., Vojta, P.J., Ropp, P.A., and Copeland, W.C. (1999). Cloning and chromosomal mapping of the human DNA polymerase theta (POLQ), the eighth human DNA polymerase. Genomics *59*, 90-96.

Sheaff, R., Ilsley, D., and Kuchta, R. (1991). Mechanism of DNA polymerase alpha inhibition by aphidicolin. Biochemistry *30*, 8590-8597.

Shima, N., Alcaraz, A., Liachko, I., Buske, T.R., Andrews, C.A., Munroe, R.J., Hartford, S.A., Tye, B.K., and Schimenti, J.C. (2007). A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice. Nat Genet *39*, 93-98.

Shima, N., Hartford, S.A., Duffy, T., Wilson, L.A., Schimenti, K.J., and Schimenti, J.C. (2003). Phenotype-based identification of mouse chromosome instability mutants. Genetics *163*, 1031-1040.

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

Simsek, D., Brunet, E., Wong, S.Y., Katyal, S., Gao, Y., McKinnon, P.J., Lou, J., Zhang, L., Li, J., Rebar, E.J., Gregory, P.D., Holmes, M.C., and Jasin, M. (2011a). DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation. PLOS Genet *7*, e1002080.

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

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

Sobol, R.W., Horton, J.K., Kuhn, R., Gu, H., Singhal, R.K., Prasad, R., Rajewsky, K., and Wilson, S.H. (1996). Requirement of mammalian DNA polymerase-beta in base-excision repair. Nature *379*, 183-186.

Sotiriou, C., Wirapati, P., Loi, S., Harris, A., Fox, S., Smeds, J., Nordgren, H., Farmer, P., Praz, V., Haibe-Kains, B., Desmedt, C., Larsimont, D., Cardoso, F., Peterse, H., Nuyten, D., Buyse, M., Van de Vijver, M.J., Bergh, J., Piccart, M., and Delorenzi, M. (2006). Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst *98*, 262-272.

Stavnezer, J., Guikema, J.E., and Schrader, C.E. (2008). Mechanism and regulation of class switch recombination. Annu Rev Immunol *26*, 261-292.

Strauss, B., Rabkin, S., Sagher, D., and Moore, P. (1982). The role of DNA polymerase in base substitution mutagenesis on non-instructional templates. Biochimie *64*, 829-838.

Strauss, B.S. (2002). The "A" rule revisited: polymerases as determinants of mutational specificity. DNA Repair *1*, 125-135.

Strumberg, D., Pilon, A.A., Smith, M., Hickey, R., Malkas, L., and Pommier, Y. (2000). Conversion of topoisomerase I cleavage complexes on the leading strand of ribosomal DNA into 5'-phosphorylated DNA double-strand breaks by replication runoff. Mol Cell Biol *20*, 3977-3987.

Symington, L.S., and Gautier, J. (2011). Double-strand break end resection and repair pathway choice. Annu Rev Genet *45*, 247-271.

Takata, K., Arana, M.E., Seki, M., Kunkel, T.A., and Wood, R.D. (2010). Evolutionary conservation of residues in vertebrate DNA polymerase N conferring low fidelity and bypass activity. Nucleic Acids Res *38*, 3233-3244.

Takata, K., Shimizu, T., Iwai, S., and Wood, R.D. (2006). Human DNA polymerase N (POLN) is a low fidelity enzyme capable of error-free bypass of 5S-thymine glycol. J Biol Chem *281*, 23445-23455.

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

Tomida, J., Takata, K., Lange, S.S., Schibler, A.C., Yousefzadeh, M.J., Bhetawal, S., Dent, S.Y., and Wood, R.D. (2015). REV7 is essential for DNA damage tolerance via two REV3L binding sites in mammalian DNA polymerase zeta. Nucleic Acids Res *43*, 1000-1011.

Ukai, A., Maruyama, T., Mochizuki, S., Ouchida, R., Masuda, K., Kawamura, K., Tagawa, M., Kinoshita, K., Sakamoto, A., Tokuhisa, T., and J, O.W. (2006). Role of DNA polymerase theta in tolerance of endogenous and exogenous DNA damage in mouse B cells. Genes Cells *11*, 111-121.

Vaisman, A., Frank, E.G., McDonald, J.P., Tissier, A., and Woodgate, R. (2002). poliota-dependent lesion bypass in vitro. Mutat Res *510*, 9-22.

Vaisman, A., Lehmann, A.R., and Woodgate, R. (2004). DNA polymerases eta and iota. Adv Protein Chem *69*, 205-228.

Wang, Y., Klijn, J.G., Zhang, Y., Sieuwerts, A.M., Look, M.P., Yang, F., Talantov, D., Timmermans, M., Meijer-van Gelder, M.E., Yu, J., Jatkoe, T., Berns, E.M., Atkins, D., and Foekens, J.A. (2005). Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet *365*, 671-679.

Ward, I.M., Minn, K., Jorda, K.G., and Chen, J. (2003). Accumulation of checkpoint protein 53BP1 at DNA breaks involves its binding to phosphorylated histone H2AX. J Biol Chem *278*, 19579-19582.

Weill, J.C., and Reynaud, C.A. (2008). DNA polymerases in adaptive immunity. Nat Rev Immunol *8*, 302-312.

Weinstock, D.M., Elliott, B., and Jasin, M. (2006). A model of oncogenic rearrangements: differences between chromosomal translocation mechanisms and simple double-strand break repair. Blood *107*, 777-780.

Welsh, C., Day, R., McGurk, C., Masters, J.R., Wood, R.D., and Koberle, B. (2004). Reduced levels of XPA, ERCC1 and XPF DNA repair proteins in testis tumor cell lines. Int J Cancer *110*, 352-361.

White, T.B., and Lambowitz, A.M. (2012). The retrohoming of linear group II intron RNAs in Drosophila melanogaster occurs by both DNA ligase 4-dependent and -independent mechanisms. PLOS Genet *8*, e1002534.

Wilson, P., Liu, Y.J., Banchereau, J., Capra, J.D., and Pascual, V. (1998). Amino acid insertions and deletions contribute to diversify the human Ig repertoire. Immunol Rev *162*, 143-151.

Woodman, I.L., and Bolt, E.L. (2009). Molecular biology of Hel308 helicase in archaea. Biochem Soc Trans *37*, 74-78.

Yamanaka, K., Minko, I.G., Takata, K., Kolbanovskiy, A., Kozekov, I.D., Wood, R.D., Rizzo, C.J., and Lloyd, R.S. (2010). Novel enzymatic function of DNA polymerase nu in translesion DNA synthesis past major groove DNA-peptide and DNA-DNA cross-links. Chemical research in toxicology *23*, 689-695.

Yamane, A., Resch, W., Kuo, N., Kuchen, S., Li, Z., Sun, H.W., Robbiani, D.F., McBride, K., Nussenzweig, M.C., and Casellas, R. (2011). Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. Nat Immunol *12*, 62-69.

Yang, Y., McBride, K.M., Hensley, S., Lu, Y., Chedin, F., and Bedford, M.T. (2014). Arginine methylation facilitates the recruitment of TOP3B to chromatin to prevent R loop accumulation. Mol Cell *53*, 484-497.

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

Yoshimura, M., Kohzaki, M., Nakamura, J., Asagoshi, K., Sonoda, E., Hou, E., Prasad, R., Wilson, S.H., Tano, K., Yasui, A., Lan, L., Seki, M., Wood, R.D., Arakawa, H., Buerstedde, J.M., Hochegger, H., Okada, T., Hiraoka, M., and Takeda, S. (2006). Vertebrate POLQ and POLbeta cooperate in base excision repair of oxidative DNA damage. Mol Cell *24*, 115-125.

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

Yousefzadeh, M.J., Wyatt, D.W., Takata, K., Mu, Y., Hensley, S.C., Tomida, J., Bylund, G.O., Doublie, S., Johansson, E., Ramsden, D.A., McBride, K.M., and Wood, R.D. (2014). Mechanism of suppression of chromosomal instability by DNA polymerase POLQ. PLoS Genet *10*, e1004654.

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

Yu, X., Tsibane, T., McGraw, P.A., House, F.S., Keefer, C.J., Hicar, M.D., Tumpey, T.M., Pappas, C., Perrone, L.A., Martinez, O., Stevens, J., Wilson, I.A., Aguilar, P.V., Altschuler, E.L., Basler, C.F., and Crowe, J.E., Jr. (2008). Neutralizing antibodies derived from the B cells of 1918 influenza pandemic survivors. Nature *455*, 532-536.

Zahn, K.E., Averill, A.M., Aller, P., Wood, R.D., and Doublié, S. (2015). Human DNA polymerase theta grasps the primer temrinus to mediate DNA repair. Nat Struct Mol Biol. doi:10.138/nsmb2993.

Zan, H., Shima, N., Xu, Z., Al-Qahtani, A., Evinger Iii, A.J., Zhong, Y., Schimenti, J.C., and Casali, P. (2005). The translesion DNA polymerase theta plays a dominant role in immunoglobulin gene somatic hypermutation. EMBO J *24*, 3757-3769.

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

Zhang, Y., Yuan, F., Wu, X., Taylor, J.S., and Wang, Z. (2001). Response of human DNA polymerase iota to DNA lesions. Nucleic Acids Res *29*, 928-935.

Zietlow, L., Smith, L.A., Bessho, M., and Bessho, T. (2009). Evidence for the involvement of human DNA polymerase N in the repair of DNA interstrand cross-links. Biochemistry *48*, 11817-11824.

Vita

Matthew James Yousefzadeh was born in Charleston, South Carolina, in 1982 to Rachel and Mohammed Yousefzadeh. He attended the College of Charleston, where he obtained his Bachelors of Science in Biology in 2005. For the next two years he worked as a Laboratory Specialist II and adjunct instructor at the College of Charleston. Then he attended the Medical University of South Carolina, where he obtained his Masters of Science in Cytology and Bioscience in 2008. He then worked as a Research Specialist II at the Medical University of South Carolina before joining the Graduate School of Biomedical Sciences at The University of Texas M.D. Anderson Cancer Center, in August 2008.